BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

REMARKS

In view of the above amendments and the following remarks, reconsideration of the outstanding office action is respectfully requested.

The objection to the specification is obviated in view of the above amendments.

The rejection of claims 1, 3-8, 13-20, and 22 under 35 U.S.C. § 112, first paragraph, for lack of written description is respectfully traversed in view of the above amendments and the following remarks. Support for the amendments to claim 1 is found, for example, in original claims 1-2, as well as page 7, lines 28-33, page 13, line 21 to page 14, line 5, page 24, line 6 to page 27, line 6, page 36, lines 6-7, and page 40, line 16 to page 61, line 5 of the specification.

It is the position of the U.S. Patent and Trademark Office ("PTO") that the specification does not disclose identifying functional groups which bind to catalytic residues of a protein kinase, covalently attaching the first module to a peptide scaffold, and attaching the second module to the identified first module by substituting the second module for the peptide scaffold. Moreover, the PTO argues that the specification does not sufficiently teach whether the first module contains a peptide scaffold and the functional groups are present on the first module or the first module is distinct from the peptide scaffold. In addition, the PTO asserts that the specification does not recite that identifying the first module comprises attaching the first module to a peptide scaffold. Applicants respectfully disagree.

In particular, applicants respectfully direct the PTO's attention to Figures 1-3 of the present application and the description at page 11, line 1 to page 22, line 7 of the specification, where a detailed description and specific examples of the identification of functional groups which bind to the conserved catalytic residues of a protein kinase is set forth. More specifically, the specification discloses the use of initial molecular modeling studies to model candidate first module (M_1) functional groups in the conserved catalytic region of the serine kinase cAMP-dependent protein kinase ("PKA") active site (see Figure 3 and page 11, line 1 to page 13, line 9) and the subsequent formation of specific pentapeptide-based inhibitors which include an M_1 functional group covalently bound to a pentapeptide sequence, based on initial modeling studies for PKA and $pp60^{c\text{-}src}$ (page 13, line 10 to page 22, line 7 of the specification). The specification discloses the use of two different pentapeptide scaffolds including: (1) Ac-Arg-Arg-Gly-Xaa-Ile-NH_2 (see Table I), and (2) Ac-Ile-Xaa-Gly-Glu-Phe-NH_2 (see Table II). The Xaa in the first sequence is Ala covalently

bonded to an M_1. The Xaa in the second sequence is Phe covalently bonded to an M_1. As shown in Table I, eleven (11) different functional groups for M_1 were tested and as shown in Table II, eight (8) different functional groups for M_1 were tested, including, but not limited to, phosphonic acid, sulfamic acid, carboxylic acid, aldehyde, and amide functional groups. The pentapeptide-based inhibitors, which include an M_1 functional group covalently bound to a pentapeptide sequence, were then tested in two different assays (Literature Mimetic assay conditions (L) and Cellular Mimetic assay conditions (C)) and suitable M_1 functional groups were identified with protein kinase inhibitory activity (page 13, line 10 to page 17, line 21). In addition, as shown in Table III and the accompanying description, four (4) different boronic acid functional groups for M_1 were tested under the above two assay conditions and suitable M_1 functional groups were identified (page 17, line 22 to page 22, line 7). Accordingly, the specification discloses identifying functional groups which bind to catalytic residues of a protein kinase (i.e., show protein kinase inhibitory activity) for PKA and $pp60^{c\text{-}src}$ and covalently attaching the first module to a peptide scaffold. Specific structures for functional groups are set forth, as well as a disclosed correlation between their function (binding to catalytic residues of a protein kinase) and structure, based on molecular modeling studies and production and testing of pentapeptide-based inhibitors (page 11, line 1 to page 22, line 7 of the specification). In addition, specific methods for covalently attaching the first module to a peptide scaffold are set forth at page 14, lines 2-5 and page 18, lines 24-27 (references are incorporated by reference at page 65, lines 3-4). Moreover, additional peptide scaffolds for specific protein kinase inhibitors which can be used as starting materials in the method of the present invention are known in the art and are described, for example, in Pearson et al., "Protein Kinase Phosphorylation Site Sequences and Consensus Specificity Motifs: Tabulations," Methods in Enzymology 200:62-81 (1991) (copy attached at Exhibit 1); Kemp et al., "Design and Use of Peptide Substrates for Protein Kinases," Methods in Enzymology 200:121-134 (1991) (copy attached at Exhibit 2); Kemp et al., "Protein Kinase Recognition Sequence Motifs," Trends in Biochemical Sciences 15(9):342-346 (1990) (copy attached at Exhibit 3); Sparks et al., "Molecular Basis for Substrate Specificity of Protein Kinases and Phosphatases," Intl. J. Biochem. 18(6):497-504 (1986) (copy attached at Exhibit 4); Ruzzene et al., "Assay of Protein Kinases and Phosphatases Using Specific Peptide Substrates," Protein Phosphorylation, 2nd Ed., Ed., Hardie, Padua, Italy, pp. 221-253 (1999) (copy attached at Exhibit 5); Tegge et al., "Analysis of Protein Kinase Substrate Specificity by the Use of Peptide Libraries on Cellulose Paper (SPOT-Method)," Methods in Molecular Biology 87:99-106 (1998) (copy attached at Exhibit 6); Zhou et al., "The Use of

Peptide Library for the Determination of Kinase Peptide Substrates," Methods in Molecular Biology 87:87-98 (1998) (copy attached at Exhibit 7); Engstroem et al., "Detection and Identification of Substrates for Protein Kinases: Use of Proteins and Synthetic Peptides," Methods in Enzymology 107:130-54 (1984) (copy attached at Exhibit 8); Casnellie et al., "The Use of Synthetic Peptides for Defining the Specificity of Tyrosine Protein Kinases," Advances in Enzyme Regulation 22:501-15 (1984) (copy attached at Exhibit 9); and Fukunaga et al., "Identifying Protein Kinase Substrates by Expression Screening with Solid-Phase Phosphorylation," Protein Phosphorylation, 2nd Ed., Ed., Hardie, Padua, Italy, pp. 291-313 (1999) (copy attached at Exhibit 10). Therefore, those of ordinary skill in the art, using the methods and structures disclosed in the present application, would have been able to identify functional groups which bind to catalytic residues of a protein kinase (in addition to PKA or $pp60^{s\text{-}src}$) and covalently attach a first module to a peptide scaffold.

With regard to the PTO's position that the specification does not disclose attaching the second module to the identified first module by substituting the second module for the peptide scaffold, this rejection is respectfully traversed in view of the above amendments to claim 1.

With regard to the PTO's position that the specification does not sufficiently teach whether the first module contains a peptide scaffold and the functional groups are present on the first module or the first module is distinct from the peptide scaffold, applicants respectfully direct the PTO's attention to Figure 1 of the above-identified application. In Figure 1 and the accompanying description, it is clearly set forth that the first module is distinct and different from the peptide scaffold, but is initially bound to the peptide scaffold.

With regard to the PTO's position that the specification does not recite that identifying the first module comprises attaching the first module to a peptide scaffold, this rejection is respectfully traversed in view of the above amendments to claim 1.

Accordingly, the rejection of claims 1, 3-8, 13-20, and 22 for lack of written description is improper and should be withdrawn.

The rejection of claims 1, 3-8, 13-20, and 22 under 35 U.S.C. § 112, first paragraph, for lack of enablement is respectfully traversed.

It is the PTO's position that the claims omit essential structures for the first module, peptide scaffold, and non-peptide scaffold and structural relationships of the reagents. Moreover, the PTO asserts that undue experimentation would be required to make and use the invention since the specification fails to give adequate guidance on how to identify a first module and use the first module covalently attached to a peptide scaffold to attach to a

second module by substituting the second module for a peptide scaffold. The PTO further asserts that the specification does not recite any compounds in which the second module is substituted for a peptide scaffold and use of these compounds as protein kinase inhibitors. In addition, the PTO argues that the breadth of the claims is open-ended regarding the resulting protein kinase inhibitor, the state of the prior art indicates that protein kinase inhibitors are specific to the kinases and in general known to be difficult, resulting in non-functional compounds, and the art is unpredictable because organic synthesis and screening for active compounds is unpredictable when applied to compounds of diversity. Applicants respectfully disagree.

With regard to the PTO's position that the claims omit the essential structures, such as the first module, peptide scaffold, and non-peptide scaffold, applicants assert that the claims, as amended, set forth the essential structures for practicing the claimed invention and fully enable one of ordinary skill in the art to practice the claimed invention. In particular, the claims recite a first module, wherein the (at least one) first module is covalently attached to a peptide scaffold. As set forth in amended claim 1 and in the specification, for example, at page 13, line 31 to page 14, line 5, the peptide scaffold is a peptide substrate for a protein kinase of interest, wherein the length of the peptide scaffold is generally the minimum number of amino acids to give good binding to the kinase (typically and in the examples in the specification, five amino acids). Moreover, as set forth in claim 1, the first module includes one or more functional groups each of which is capable of covalently or non-covalently binding to catalytic residues of the protein kinase. Thus, the one or more functional groups of the first module must have steric and electronic characteristics that are complementary to the catalytic residues that are focused in this region of the active site. Accordingly, the functional groups of the first module are positioned to interact with the catalytic resides of the protein kinase. As set forth in the specification, initial candidate first modules comprising one or more functional groups are obtained from molecular modeling studies with kinase crystal structures, but are experimentally tested while attached to a peptide scaffold (see, e.g., page 11, line 1 to page 22, line 7 of the specification).

As set forth in the May 23, 2003, Amendment, the second module comprises a non-peptide scaffold, which includes molecules with amide or peptide bonds, so long as a part of the molecule is not a peptide. As set forth in the claims, the second module is substituted for (i.e., replaces) the peptide scaffold, which is a substrate for the protein kinase. As such, the second module must have steric and electronic characteristics that mimic the peptide substrate at the phosphorylatable position and adjacent amino acid binding sites (i.e.,

is capable of occupying the same binding region of the protein kinase as the peptide scaffold). As set forth in the specification, kinase crystal structures are used to design candidate second modules that are subsequently tested experimentally (see page 22, line 8 to page 27, line 6 of the specification). For example, indole and naphthalene second modules replace the tyrosine residue and adjacent amino acids in the $pp60^{s\text{-}src}$ peptide substrate in the Examples of the present application.

With regard to the PTO's position that the claims omit the structural relationships of the reagents, applicants respectfully disagree. In particular, the structural relationships between the first module, the second module, the functional groups of the first module, the protein kinase, and the peptide scaffold are clearly set forth in the claims and in the specification.

More specifically, amended claim 1 recites "[a] method for identifying inhibitors of protein kinases comprising: covalently attaching at least one first module to a peptide scaffold which comprises a peptide substrate for a protein kinase and identifying one or more functional groups on the first module each of which is capable of covalently or non-covalently binding to catalytic residues of the protein kinase; producing one or more combinations of the at least one first module covalently attached to at least one second module which comprises a non-peptide scaffold, wherein the at least one second module comprises an indole, and wherein the at least one second module is substituted for the peptide scaffold and is capable of occupying the same binding region of the protein kinase as the peptide scaffold; screening the one or more combinations of the first and second modules for protein kinase inhibition; and selecting combinations of the first and second modules which inhibit protein kinase activity." Thus, claim 1 clearly sets forth that the first module is initially covalently attached to a peptide scaffold and one or more complexes (or combinations) are subsequently produced including the first module covalently attached to the second module, such that the second module is substituted for the peptide scaffold. In addition, claim 1 sets forth that the first module includes one or more functional groups each of which is capable of covalently or non-covalently binding to catalytic residues of a protein kinase. Thus, the structural cooperative relationships between the first module, the second module, the functional groups of the first module, the protein kinase, and the peptide scaffold are clearly set forth in amended claim 1.

With regard to the PTO's position that the specification fails to give adequate guidance on how to identify a first module, applicants respectfully disagree. As described above, a detailed description and specific examples of the identification of first modules (M_1)

is set forth in Figures 1-3 of the present application and the description at page 11, line 1 to page 22, line 7 of the specification. More specifically, the specification discloses the use of initial molecular modeling studies to model candidate M_1 functional groups in the conserved catalytic region of the PKA active site (see Figure 3 and page 11, line 1 to page 13, line 9) and the subsequent formation of specific pentapeptide-based inhibitors which include an M_1 functional group covalently bound to a pentapeptide sequence, based on these initial modeling studies (page 13, line 10 to page 22, line 7 of the specification). The specification discloses the use of two different pentapeptide scaffolds including: (1) Ac-Arg-Arg-Gly-Xaa-Ile-NH_2 (see Tables I and III), and (2) Ac-Ile-Xaa-Gly-Glu-Phe-NH_2 (see Table II). The Xaa in the first sequence is Ala covalently bonded to an M_1. The Xaa in the second sequence is Phe covalently bonded to an M_1. As shown in Tables I-III, 23 different functional groups for M_1 were tested including, but not limited to, phosphonic acid, sulfamic acid, carboxylic acid, aldehyde, amide, and boronic acid functional groups. The pentapeptide-based inhibitors, which include an M_1 functional group covalently bound to a pentapeptide sequence, were then tested in two different assays (Literature Mimetic assay conditions (L) and Cellular Mimetic assay conditions (C)) and suitable M_1 functional groups were identified (page 13, line 10 to page 22, line 7). Accordingly, the specification discloses identifying first modules which comprise functional groups which bind to catalytic residues of a protein kinase for PKA and $pp60^{c\text{-}src}$. Specific structures for functional groups of the first module are set forth, as are methods for covalently attaching M_1s including one or more functional groups to peptide scaffolds (page 14, lines 2-5 and page 18, lines 24-27) and specific assay conditions for testing potential functional groups of a first module (page 13, line 10 to page 22, line 7). Moreover, additional peptide scaffolds for specific protein kinase inhibitors which can be used as starting materials in the method of the present invention are known in the art and are described, for example, in Pearson et al., "Protein Kinase Phosphorylation Site Sequences and Consensus Specificity Motifs: Tabulations," Methods in Enzymology 200:62-81 (1991); Kemp et al., "Design and Use of Peptide Substrates for Protein Kinases," Methods in Enzymology 200:121-134 (1991); Kemp et al., "Protein Kinase Recognition Sequence Motifs," Trends in Biochemical Sciences 15(9):342-346 (1990); Sparks et al., "Molecular Basis for Substrate Specificity of Protein Kinases and Phosphatases," Intl. J. Biochem. 18(6):497-504 (1986); Ruzzene et al., "Assay of Protein Kinases and Phosphatases Using Specific Peptide Substrates," Protein Phosphorylation, 2nd Ed., Ed., Hardie, Padua, Italy, pp. 221-253 (1999); Tegge et al., "Analysis of Protein Kinase Substrate Specificity by the Use of Peptide Libraries on Cellulose Paper (SPOT-Method)," Methods in Molecular Biology

87:99-106 (1998); Zhou et al., "The Use of Peptide Library for the Determination of Kinase Peptide Substrates," Methods in Molecular Biology 87:87-98 (1998); Engstroem et al., "Detection and Identification of Substrates for Protein Kinases: Use of Proteins and Synthetic Peptides," Methods in Enzymology 107:130-54 (1984); Casnellie et al., "The Use of Synthetic Peptides for Defining the Specificity of Tyrosine Protein Kinases," Advances in Enzyme Regulation 22:501-15 (1984); and Fukunaga et al., "Identifying Protein Kinase Substrates by Expression Screening with Solid-Phase Phosphorylation," Protein Phosphorylation, 2nd Ed., Ed., Hardie, Padua, Italy, pp. 291-313 (1999) (copies attached as Exhibits 1-10, respectively, as discussed *supra*). Therefore, those of ordinary skill in the art, using the methods and structures disclosed in the present application, would have been able to identify first modules including functional groups which bind to catalytic residues of a protein kinase (in addition to PKA or $pp60^{s\text{-}src}$) and covalently attach a first module to a peptide scaffold without undue experimentation.

With regard to the PTO's position that the specification fails to give adequate guidance on how to use the first module covalently attached to a peptide scaffold to attach to a second module by substituting the second module for the peptide scaffold and that the specification doesn't recite any compounds in which the second module is substituted for a peptide scaffold and use of the compounds as protein kinase inhibitors, this rejection is respectfully traversed in view of the above amendments to claim 1 and the following remarks.

In particular, claim 1, as amended, recites "producing one or more combinations of the at least one first module covalently attached to at least one second module . . . wherein the at least one second module is substituted for the peptide scaffold." Support for this limitation is found, for example, at page 36, lines 6-7 and page 22, line 8 to page 27, line 6 of the specification. In particular, at page 22, line 8 to page 25, line 13 of the specification, at least one second module (i.e., naphthalene, isoquinoline, or indole) is substituted for a peptide scaffold in molecular modeling studies using the first modules (e.g., boronic acid, phosphonate, and sulfamic acid) previously identified. Subsequently, combinations of the at least one first module covalently attached to at least one second module are produced and tested for protein kinase inhibition (page 25, line 14 to page 27, line 6 of the specification). Moreover, detailed examples and procedures for producing protein kinase inhibitors in which a second module, e.g., naphthalene or indole, is substituted for a peptide scaffold are set forth at page 40, line 16 to page 61, line 5. In particular, in Example 1 of the specification (page 40, line 16 to page 46, line 23), combinations of a first module (i.e., OH) covalently attached to at least one second module (i.e., naphthalene) are produced,

wherein the second module (i.e., naphthalene) is substituted for a peptide scaffold in molecular modeling studies (page 40, lines 16-28). Moreover, in Examples 2-4 (page 46, line 25 to page 61, line 5) of the specification, combinations of a first module (e.g., OH, boronic acid, and phosphonic acid) covalently attached to at least one second module (i.e., indole) are produced, wherein the second module (i.e., indole) is substituted for a peptide scaffold used in molecular modeling studies (page 47, lines 5-15). Thus, in view the amount of direction and representative examples provided in the specification and the knowledge in the art of peptide substrates for protein kinases which can be used to identify suitable first modules using the methodology described in the specification (see above), it would not require undue experimentation to produce one or more combinations of the at least one first module covalently attached to at least one second module . . . wherein the at least one second module is substituted for the peptide scaffold, as required by amended claim 1.

With regard to the PTO's position that the breadth of the claims is open-ended regarding the resulting protein kinase inhibitor, applicants assert that the breadth of the claims is fully enabled by the disclosure of the present application in view of the previous remarks.

Accordingly, the rejection of claims 1, 3-8, 13-20, and 22 for lack of enablement is improper and should be withdrawn.

The rejection of claims 1, 3-8, 13-20, and 22 under 35 U.S.C. § 112, first paragraph, for lack of written description is respectfully traversed.

It is the PTO's position that the newly added limitations "wherein said identifying one or more functional groups on the first module which preferentially bind to catalytic residues of the protein kinase" and "wherein said covalently attaching comprises substituting the at least one second module for the peptide scaffold" recited in claim 1 have no clear support in the specification and the claims as filed. Thus, the PTO asserts that this is new matter. Applicants respectfully disagree.

In particular, original claim 2, as filed, provides support for both limitations recited in claim 1. Accordingly, the rejection of claims 1, 3-8, 13-20, and 22 for new matter is improper and should be withdrawn.

The rejection of claims 1, 3-8, 13-20, and 22 under 35 U.S.C. § 112, second paragraph, for indefiniteness is respectfully traversed in view of the above amendments and the following remarks.

The PTO requests clarification of the definition of the first module. As set forth in the Amendment submitted on May 23, 2002, a module, i.e., first module, is defined to include a single molecular entity or a collection of functional groups. Support for this

definition is found at page 36, lines 8-11 of the specification, which sets forth that "[p]referred first modules have a functional group . . . The compounds of the present invention may have two or more functional groups within the first module. More preferred modules are boronic acid groups, a hydroxyl group, or an amide group." In addition, in Figure 1 the first module (i.e., M_1) is defined as "a selection of validated functionalities for binding to the conserved catalytic residues." Also, as shown in Figure 1, the first module is distinct and different from the peptide scaffold, but is initially bound to the peptide scaffold.

In addition, the PTO states that it is not clear whether the protein kinase to which the functional groups of the first module preferentially bind is the same as the peptide scaffold to which the first module is covalently attached. Applicants respectfully disagree.

Definiteness of claim language must be analyzed in view of "(A) The content of the particular application disclosure; (B) The teachings of the prior art; and (C) The claim interpretation that would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made." MPEP, § 2173.02.

The specification of the above-identified application clearly sets forth (e.g., in Figure 1) that functional groups of the first module bind to the catalytic residues (e.g., conserved catalytic site in Figure 1) of a protein kinase which is separate and distinct from the peptide scaffold to which the first module is covalently attached. Thus, the first step of the claimed method comprises producing a complex including a first module having one or more functional groups bound to a peptide scaffold, wherein the one or more functional groups of the first module are each capable of binding with the catalytic residues of a protein kinase.

The PTO further asserts that it is not clear what is the resulting end product of covalently attaching the first module to a second module by substituting the second module's non-peptide scaffold for a peptide scaffold and requests clarification. Again, as set forth in the May 23, 2002, Amendment, the second module comprises a non-peptide scaffold, which includes molecules with amide or peptide bonds, so long as a part of the molecule is not a peptide. Thus, the resulting end product of the second step of the claimed method is the production of a complex including a first module (which includes one or more functional groups) and a second module (which comprises a non-peptide scaffold), wherein the second module includes an indole as at least part of its structure and is substituted for the peptide scaffold in the complex produced in the first step of the claimed method.

In addition, the PTO argues that claim 1 omits essential structural cooperative relationships between the first module, the second module, the functional groups of the first

module, the protein kinase, and the peptide scaffold. Applicants respectfully disagree in view of the previous remarks.

The remaining rejections under 35 U.S.C. § 112, second paragraph, are respectfully traversed in view of the above amendments.

In view of the all of the foregoing, applicants submit that this case is in condition for allowance and such allowance is earnestly solicited.

Respectfully submitted,

Date: April 8, 2004

Michael L. Goldman
Registration No. 30,087

NIXON PEABODY LLP
Clinton Square, P.O. Box 31051
Rochester, New York 14603
Telephone: (585) 263-1304
Facsimile: (585) 263-1600

CERTIFICATE OF MAILING OR TRANSMISSION [37 CFR 1.8(a)]

I hereby certify that this correspondence is being:

☒ deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Mail Stop RCE, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450

☐ transmitted by facsimile on the date shown below to the United States Patent and Trademark Office at (703) _____.

April 8, 2004
Date

Signature

Wendy L. Barry
Type or Print Name

EXHIBIT 1

smaller trees from which a composite tree that includes all 117 sequences was drawn (Fig. 4). These smaller trees, each containing about 40–45 sequences, were calculated in approximately three central processing unit hours on a VAX-6220 equipped with 64 megabytes of memory running the virtual memory operating system (VMS).

The potential usefulness of the catalytic domain phylogenetic tree as a tool for classification is illustrated by the clustering of protein kinases that have similar properties. This clustering can be used to define protein kinase subfamilies. The densest cluster is composed of the 42 known or putative protein-tyrosine kinases (Fig. 4A and B). This large subfamily includes both cytoplasmic enzymes (Src, Abl, Fes, etc.) as well as the receptors for growth and differentiation factors. Two distinct protein-serine/threonine kinase clusters are recognized as containing members whose activities are under similar modes of regulation. These are the cyclic nucleotide-dependent protein kinase subfamily (13 members) and the protein kinase C subfamily (10 members). They map quite near one another within the tree (Fig. 4A and C). In addition, the protein-serine/threonine kinases regulated by Ca^{2+}/calmodulin map near one another and fall within a broad cluster. It is uncertain, however, if all members of this cluster (e.g., RSK-1C, RSK-2C, PSK-H1) are regulated by Ca^{2+}/calmodulin.

Clearly, some structure–function relationships have been maintained during the course of catalytic domain evolution. We predict other such relationships will emerge with the continued identification and functional characterization of novel protein kinases.

Acknowledgments

S.K.H. was supported by United States Public Health Service Grant GM-38793. We thank Tony Hunter for continued encouragement and Lisa Caballero for many useful comments.

[3] Protein Kinase Phosphorylation Site Sequences and Consensus Specificity Motifs: Tabulations

By RICHARD B. PEARSON and BRUCE E. KEMP

Introduction

The phosphorylation site sequences for protein-serine/threonine and protein-tyrosine kinases are presented in Table I of this chapter. Over 240 phosphorylation site sequences are included with the phosphorylated residue(s) indicated by an asterisk. Plant and prokaryote phosphorylation

site sequences have not been included. References are list d at the end of the chapter. Anyone wanting an annual update of this tabulation and/or wanting to contribute a new sequence should contact the authors.

Note that residue numbers followed by a # in the tabulation have been obtained from SWISSPROT protein database. (P) denotes a phosphorylated residue which acts as a substrate-specificity determinant. The question marks indicate potential sites within a phosphorylated peptide or sites inferred from mutagenesis studies or the stoichiometry of phosphorylation (see Table I, footnotes *c* and *h*, respectively).

Table II contains consensus phosphorylation site motifs for each enzyme, where applicable. The frequencies listed are derived from Table I unless indicated otherwise. The *S*:*T* ratio is for the total number of phosphorylation sites. Asterisks indicate the phosphorylated residue and specificity determinants are shown in boldface type.

The one-letter code for amino acids in Tables I and II is as follows: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

TABLE I
PROTEIN KINASE PHOSPHORYLATION SITE SEQUENCES[a]

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
Protein-serine/threonine kinases			
AMP-activated protein kinase (previously called acetyl-CoA carboxylase kinase-3, HMG-CoA reductase kinase, hormone-sensitive lipase kinase)	M_{74} R S S M S* G L H L	Acetyl-CoA carboxylase (rat)	42
	L_{1196} N R M S* F A S N	Acetyl-CoA carboxylase (rat)	42
	G_{1209} M TH V A S* V S D V L L D	Acetyl-CoA carboxylase (rat)	49
	M_{560} R R S V S* E A A L	Hormone-sensitive lipase (rat)	42, 49, 91
	H_{865} M V H N R S* K I N L Q D L	HMG-CoA reductase (rat)	49
β-Adrenergic receptor kinase[c]	G_{353} Y S[?]* S[?]* N G N T[?]* G E Q S[?]*	β-Adrenergic receptor[d]	50
	E_{379} D L P G T[?]* E D	β-Adrenergic receptor[d]	50
	G_{392} G T[?]* V P S[?]* D N I D S[?]* Q G R N C S[?]* T[?]* N D S[?]* L L-COOH	β-Adrenergic receptor[d]	50

(continued)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
Branched chain α-ketoacid dehydrogenase kinase	T_{326} Y R I G H H S* T S D D S S	Branched chain α-ketoacid dehydrogenase site 1[#]	45
	A_{340} Y R S* V D E V N Y W D K	Branched chain α-ketoacid dehydrogenase site 2[#]	45
Calmodulin-dependent protein kinase I	N_2 Y L R R R L S* D S N F M	Synapsin I site 1	14, 100
Calmodulin-dependent protein kinase II	T_{564} R Q T S* V S G Q A P P K	Synapsin I site 2 (bovine)	14, 100
	T_{601} R Q A S* Q A G P M P R	Synapsin I site 3 (bovine)	14, 100
	R_{13} R A S T* I E M P Q Q A R	Phospholamban	15, 78
	P_1 L S R T L S* V S S	Glycogen synthase site 2	16
	K_{191} M A R V F S* V L R E	Calcineurin[#]	17, 139
	R_{15} R A V S* E Q D A K	Tyrosine hydroxylase (monooxygenase) site C	18
	R_{37} R Q S* L I E D A R K	Tyrosine hydroxylase (monooxygenase) site A	18
	R_{13} K L S* D/N F G E/Q	Phenylalanine hydroxylase (monooxygenase)	19
	Y_7 L R R A S* V A Q L T* Q E	Pyruvate kinase[#]	20
	R_{10} S K Y L A S* A S T M	Myelin basic protein	21
	R_{64} T T H Y G S* L P Q K	Myelin basic protein	21
	K_{91} N I V T* P R T P P	Myelin basic protein	21
	R_{113} F S* W G A E G Q K	Myelin basic protein	21
	R/K T A S* F S E S R	ATP-citrate lyase	43
	F_{19} I I G S V S* E D N	Acetyl-CoA carboxylase[#]	28
	M_{560} R R S V S* E A A L	Hormone-sensitive lipase	70
	R_{808} A I G R L S S* M A M	Smooth muscle myosin light chain kinase	130, 131
	R_{283} Q E T* V D C L K K F N A R R K L K	Autophosphorylation (α subunit)	96–98
	R_{284} Q E T* V E C L K K F N A R R K L K	Autophosphorylation (β subunit)	98, 99

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
Calmodulin-dependent protein kinase III	A_{51}G E T[?]* R F T[?]* D T[?]* R	Elongation factor 2	22
Casein kinase I	V_{22} S(P) S(P) S(P) E E S* I I S	α_{s2}-Casein[#]	29
	L_{143} S(P) T S(P) E E N S* K K	α_{s2}-Casein[#]	29
	L_{31} S(P) S(P) S(P) E E S* I T R	β-Casein[#]	29
	V_{37} N E L S* K D I	α_{s1}-Casein[#]	29
	G_{45} S(P) E S(P) T* E D Q	α_{s1}-Casein[#]	29
	P_1 L S* R T L S*(P) V S S* L P G L	Glycogen synthase	30, 118
	K_{23} R S G S* V/I Y E P L K	Phosphorylase kinase (β subunit)[#]	102
Casein kinase II[d]	R_{40} L S* E H S* S P E E E A	DARPP-32 (bovine)	25
	E_{98} N Q A S* E E E D E L G E	DARPP-32 (bovine)	25
	E_{225} R D K E V S* D D E A E E	α-Heart shock protein-90 (α-Hsp-90)	26
	E_{225} R E K E I S* D D E A E E	β-Hsp-90 (human)	26
	E_{257} I E D V G S* D E E E E	α-Hsp 90 (human)	26
	K_{249} I E D V G S* D E E D D	β-Hsp 90 (human)	26
	G_{130} D R F T* D E E V D E	Myosin regulatory light chain, chicken smooth muscle	27
	S_{23} V S E D N S* E D E I S N L	Acetyl-CoA carboxylase[#]	28
	S[†] D E E V E	Troponin T	29
	P_{654} H Q S* E D E E E	Glycogen synthase (residue numbers based on human muscle sequence)	29, 41, 150
	A_{72} D S* E S* E D E E D	cAMP-dependent protein kinase regulatory subunit R_{II}[#]	108
	V_5 E E D A E S* E D E D E E D	Nucleolar protein B23[#]	29
	P_{85} A S* E D E D E E E D	Nucleolar protein C23[#]	29

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
	L E L S* D D D D E S K	Myosin heavy chain (brain)	56
	D_{80} D D D A Y S* D T E T T E	Phosphatase inhibitor-2	71
	E_{117} Q E S* S* G E E D S D L	Phosphatase inhibitor-2	71
	R_3 R P R H S I Y S* S* D D D E E D	c-Myb	87
	P_{245} P T?* T?* S?* S?* D S?* E E E Q E D E E E	Myc	88
	S_{344} P R S?* S?* D T E E N	Myc	88
	S_9 S S* E S* G A P E A A E E D	Clathrin light chain LC_b	109
	S^*_{89} D E E D E E	Elongation factor 1β	110
	D_{78} T* D S E E E I R E	Calmodulin	116
	D_{95} G D G Y I S* A A E L R H	Calmodulin	116
	E_{26} Q L N D S* S* E E E D E I D	Human papillomavirus E7 oncoprotein (HPV E7)	117
	T_{301} G S* D D E D E S N E Q	Ornithine decarboxylase[#]	29
	E_{84} E S P A S* D E A E E K	High mobility group 14 protein	140
Crystallin kinase(s)	R_{116} R Y R L P S* N V D	α_A-Crystallin	58
	R_{12} P F F P F H S* P S R	α_B-Crystallin	59
	P_{39} A S T S L S* P F Y L R P P	α_B-Crystallin	59
cAMP-dependent protein kinase (mammalian)[c]	Y_7 L R R A S* L/V A Q L T	Pyruvate kinase[#]	1
	F_1 R R L S* I S T	Phosphorylase kinase α chain	1
	Q_{692} W P R R A S* C T S	Glycogen synthase site 1a (residue numbers based on human muscle sequence)	1, 150
	G_{706} S K R S N S* V D T	Glycogen synthase site 1b (residue numbers based on human muscle sequence)	1, 150
	R_{21} T K R S G S* V/I Y E	Phosphorylase kinase β chain[#]	1

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
	A_{29} G A R R K A S* G P P	Histone H1[#] (residue numbers from bovine sequence)	1
	K_{13} A K T R S S* R A	Histone H2A[#] (residue numbers from bovine sequence)	1
	G_{26} K K R K R S* R K E S* Y S	Histone H2B[#] (residue numbers from bovine sequence)	1
	E R R K S* K S G A G	cAMP regulated phosphoprotein M_r = 21,000 (ARPP-21)	55
	Y_3 L R R R L S* D S N F	Synapsin I site 1	100
	N_{19} Y R G Y S* L G N Y V	Reduced carboxymethylated maleylated (RCMM)-lysozyme	76
	R_{27} A S* F G S R G S* G S	Desmin	86
	S_{47} R T S* A V P T	Desmin	86
	P_1 L S R T L S* V S S	Glycogen synthase site 2	1
	A_{16} V R R S* D R A	Troponin I (cardiac)	69
	M_{560} R R S* V S E A A L	Hormone-sensitive lipase	42
	M_{74} R S S* M S G L H L	Acetyl-CoA carboxylase	42
	S_{12} Q R R R S* L E P P D	pp60$^{c\text{-}src}$	79
	K_{23} R K R K S?* S?* Q C L V K	c-erbA	134
	K_{11} H K R K S?* S?* Q C L V K	v-erbA	134
	R_{45} N T D G S* T D Y G I	RCMM-lysozyme	76
	R_{212} R K G T* D V	Lipocortin I (p35, calpactin II)	81
	I_{29} R R R R P T* P A T	Phosphatase inhibitor-1[#]	1
	K P R R K D T* P A L	G substrate	1
	K_{188} R V K G R T W T* L C G T	Autophosphorylation of catalytic subunit	1, 95
	R_{89} F D R R V S* V C A	Autophosphorylation of regulatory subunit R_{II}	1
Yeast	K_{218} R K Y L K K L T R R A S* F S A	ADR1	2

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
	P_7 R R D S* T E G F	Fructose-1,6-bisphosphatase	128
	Q_{141} R R T S* V S G E	Autophosphorylation of regulatory subunit	129
cGMP-dependent protein kinase	F_1 R R L S* I S T E	Phosphorylase kinase α chain	3
	D_{25} G K K R K R S* R K E S*	Histone H2B	3
	R_{92} R R R G A I S* A E V Y	cAMP-dependent protein kinase regulatory subunit R_I	3
	K K P R R K D T* P A L H	G substrate site 1	3
	Q K P R R K D T* P A L H	G substrate site 2	3
	R_{29} R R R P T* P A M L	DARPP-32[#]	3
	Q_{28} I R R R R P T* P A T L	Phosphatase inhibitor-1	3
	P_1 K R K V S* S A E G	High mobility group 14 protein	3
	K_{17} R R S A R L S* A K P A	High mobility group 14 protein	3
	M_{560} R R S* V S E A A L	Hormone-sensitive lipase	91
	A_{11} I R R A S* T I E M	Phospholamban	*f*
	V_{45} L P V P S* T H I G P	Autophosphorylation	3
	I_{53} G P R T T* R A Q G I	Autophosphorylation	3
	P_{67} Q T Y R S* F H D L R	Autophosphorylation	3
	A_{79} F R K F T* K S E R S	Autophosphorylation	3
Double-stranded DNA-activated protein kinase	P_1 E E T* Q T* Q D Q P M E	Heat-shock protein 90α (human)	37
Double-stranded RNA-activated protein kinase (p68 kinase)	I_{45} L L S E L S* R R	eIF-2α	38
Endogenous eIF-4E kinase	K_{49} N D K S* K T W Q A N L R	eIF-4E	40
Glycogen synthase kinase-3	P_{637} R P A S* V P P S* P S L S* R H S S* P Q H S(P)	Glycogen synthase sites 3a, b, c and 4 (residue numbers based on human muscle sequence)	35, 103, 150

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
	D_{68} E P S T* P Y H S M I G D D D D A Y S(P) D	Phosphatase inhibitor-2	71
	L_{39} R E A R S* R A S* T P P	cAMP-dependent protein kinase regulatory subunit R_{II}	72
	K_{36} P G F S* P Q P S* R R G S(P)	Protein phosphatase-1 G subunit[j]	36
	A_{343} P V S?* C L G E H H H C T?* P S?* P P V D H G C L	c-Myb[#] (residue numbers based on chicken sequence)	119
	E_{238} T* P P L S* P I D M E S* Q E R	c-Jun	120
Glycogen synthase kinase-4	P_1 L S R T L S* V S S	Glycogen synthase site 2	89
	M_{560} R R S V S* E A A L	Hormone-sensitive lipase	42, 90
Growth-associated H1 histone kinase (MPF, cdc2+/CDC28 protein kinases)[g]	T_{10} S E/Q P A K T* P V K	Histone H1 (calf thymus)	53, 54
	K_{130} A T G A A T* P K	Histone H1 (calf thymus)	53, 54
	K_{152} T* P K	Histone H1 (calf thymus)	53, 54
	V_{177} A K S* P K	Histone H1 (calf thymus)	53, 54
	F_{29} P A S Q T* P N K T A	pp60[c-src]	104
	P_{41} D T H R T* P S R S F	pp60[c-src]	104
	S_{67} D T V T S* P Q R A G	pp60[c-src]	104
	D_{562} A P D T* P E L L H T K	c-Abl type IV	137, 138
	S_{583} E P A V S* P L L P R	c-Abl type IV	137, 138
	H_{122} S T* P P K K K R K	Large T antigen	137
	S_{313} S S* P Q P	p53	137
	S_{16}^{*} P T R	Lamin B	137
	S_{1}^{*} S K R A K A K T* T K K	Myosin regulatory light chain, nonmuscle	137
	K R S* P K K	SW15	137
	Y S?* P T S?* P S	RNA polymerase II	105
Growth factor-regulated kinase	P_{667} L T* P S G E A	EGF receptor	34
Heme-regulated eIF-2α kinase	I_{45} L L S* E L S* R R	eIF-2α	38, 39
Histone H4 kinase I	V_{43} K R I S* G L	Histone H4	48

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
Histone H4 kinase II	Ac-S[†] G R G K G G	Histone H4	48
Insulin receptor-associated serine kinase (IRSK)	K/R_{1292} S[?]* S[?]* H C Q R	Insulin receptor	115
Isocitrate dehydrogenase kinase	T_{104} T P V G G G I R S* L N V A	Isocitrate dehydrogenase	47
Mitogen-activated protein kinase (MAP kinase)	T_{94} P R T* P P P	Myelin basic protein	107
Mitogen-activated S6 kinase (M_r 70,000)	R_{232} R L S* S* L R A S* T S K S* E S S* Q K	Ribosomal protein S6	147, 148
Myosin I heavy chain kinase	R/K A G T* T Y A L N L N K	Myosin 1A heavy chain	32
	G_{307} G A G A K K M S* T Y N V	Myosin 1B heavy chain	32
	G_{303} E Q G R G R S S* V Y S C	Myosin 1C heavy chain	32
Myosin light chain kinase Skeletal muscle	K_5 R R A A E G S S* N V F	Myosin regulatory light chain, chicken skeletal muscle	6, 80
Smooth muscle	K_{11} K R P Q R A T* S* N V F	Myosin regulatory light chain, chicken smooth muscle	5, 60
Phosphorylase kinase	D_6 Q E K R K Q I S* V R G	Phosphorylase	4
	P_1 L S R T L S* V S S	Glycogen synthase site 2	4
	R_{21} T K R S G S* V/I Y E	Autophosphorylation (β chain)[f]	1, 101
Proline-directed protein kinase	P_2 T P S A P S* P Q P K	Tyrosine hydroxylase (monooxygenase)	33
	T_{546} R P P A S* P S P Q R	Synapsin I (bovine)	100
Protease-activated kinases I and II	A_{229} K R R R L S S* L R A	Ribosomal protein S6	52
Protein kinase C[e]	Q_4 K R P S* Q R S K Y L	Myelin basic protein	7
	K_{1329} K N G R V L T* L P R S	Insulin receptor (rat)	8
	K_{1329} K N G R I L T* L P R S	Insulin receptor (human)	121, 122
	Y_{20} T R F S* L A R	Transferrin receptor	9
	P_1 L S R T L S* V S S	Glycogen synthase site 2	135
	P_{691} Q W P R R A S* C T S	Glycogen synthase site 1a (residue numbers based on human muscle sequence)	135, 150

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
	K_{37} I Q A S* F R G H I T R K K	Neuromodulin	136
	S† S* K R A K A K T* T K K R	Myosin regulatory light chain, chicken smooth muscle	10
	K_{139} K K K K R F S* F K K S* F K L S* G F S* F K K N K K	MARCKS protein	11
	G_{97} T G A S G S* F K	Histone H1	12
	F_{43} F G S* D R G	Myelin basic protein	7
	G_{148} T L S* K I F	Myelin basic protein	7
	R_{645} R R H I V R K R T* L R R L	EGF receptor	23
	R_{89} R R H I V R K R T* L R R L	erb B	126, 127
	A_{229} K R R R L S S* L R A	Ribosomal protein S6	24
	Y_{21} V Q T* V K S* S* K G G P G	Lipocortin I (p35, calpactin II)	81, 82
	S_{21} A Y G S* V K P Y T N P D	Lipocortin II (p36, calpactin I heavy chain)	83
	G_{555} K S* S* S* Y S K	Fibrinogen	84
	H_{594} E G T H S* T K R	Fibrinogen	84
	D_{354} L K L R R S?* S?* S?* V G Y	Acetylcholine receptor	85
	R_9 V S S* Y R R T F G	Desmin	86
	R_{27} A S* F G S R G S G S S* V T S R	Desmin	86
	P_{53} T L S* T F R T T R	Desmin	86
	Ac-S† L K D H L I H N V H K	Lactate dehydrogenase	123
	W_{240} Q R R Q R K S* R R T I	Interleukin-2 receptor	124
	G_2 S S K S K P K D P S* Q R R R S	$pp60^{src}$	125
	R_{75} S S* M S G L H	Acetyl-CoA carboxylase	42, 141
	R_{88} D R K K I D S* F A S N	Acetyl-CoA carboxylase	42, 141
Pyruvate dehydrogenase kinase	R_{226} Y G M G T S* V E R	E1α pyruvate dehydrogenase[#] (residue numbers based on human sequence)	51
	R_{287} Y H G H S* M S D P G V S* Y R	E1α pyruvate dehydrogenase[#] (residue numbers based on human sequence)	51

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
Rhodopsin kinase[h]	D_{330} D E A S* T* T* V S* K T?* E T?* S* Q V A P	Rhodopsin	44
S6 kinase II (M_r 92,000)	A_{229} K R R R L S?* S?* L R A	Ribosomal protein S6	31
	I_{35} G R R Q S* L I E D A	Tyrosine hydroxylase (monooxygenase)	31
	P_1 L S R T L S* V S S L P G	Glycogen synthase site 2	31
	R_{339} G R A S S* H S S	Lamin C[#] (residue numbers based on human sequence)	31
Sperm-specific histone kinase	P_1 G S* P Q K R A A S* P R K S* P K K S* P R K A S A S* P R	Histone H1 (sea urchin sperm)	57
	K_8 R S* P T K R S* P Q K G	Histone $H2B_1$ (sea urchin sperm)	57
	R_{10} K G S* P R K G S* P K R G	Histone $H2B_2$ (sea urchin sperm)	57
Tropomyosin kinase	D_{275} H A L N D M T S* I-COOH	α-Tropomyosin	46
	D_{275} N A L N D I T S* L-COOH	β-Tropomyosin	46
Unknown kinase(s)	K_{85} K R R L S* F S E T F	Interleukin-1α	144
	R_{662} E L V E P L T* P S* G E A P	EGF receptor	145
	F_{1041} L Q R Y S* S* D P T G A L	EGF receptor	142, 145
	T_{82} F P P A P G S* P E P P	Adenovirus type 5 289R E1A protein	146
	R_{215} R P T S* P V S R E C N S S T D S* C D S G	Adenovirus type 5 289R E1A protein	146
Protein-tyrosine kinase[i]			
Colony-stimulating factor 1 receptor kinase	Q_{689} D S E G D S S Y* K N I H	Autophosphorylation	132
	K_{698} N I H L E K K Y* V R R D	Autophosphorylation	132, 133
	R_{799} D I M N D S N Y* V V K G	Autophosphorylation	132, 133
EGF-receptor kinase	S_{1166} T R E N A E Y* L R V A P Q S	Autophosphorylation	61
	I_{1141} S L D N P D Y* Q Q D F F P K	Autophosphorylation	61

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
	T_{1061} F L P V P E Y* I N Q S V P K	Autophosphorylation	61
	S_{1080} V Q N P V Y* H N Q P L N	Autophosphorylation	94
	D_{979} E E D M D D V V D A D E Y* L I P Q Q	Autophosphorylation	111, 142
	D_{135} E E V D E M Y* R E A P I D K	Myosin regulatory light chain, smooth muscle	61
	D_{148} V K G N F N Y* V E F T R I L	Myosin regulatory light chain, smooth muscle	61
	I_1 D N E E Q E Y* I K T V K G S	Lipocortin I (p35, calpactin II), porcine	61, 63
	I_{14} E N E E Q E Y* V Q T V K S S	Lipocortin I (p35, calpactin II), human	81
	L_{466} A E G S A Y* E E	Phospholipase C-γ	92
	A_{767} E P D Y* G A L Y E	Phospholipase C-γ	92
	N_{779} P G F Y* V E A N	Phospholipase C-γ	92
	E_{1251} A R Y* Q Q P F E D F R	Phospholipase C-γ	92
	T_1 D V E T T Y* A D F I A S G	Protein kinase inhibitor protein	61, 73
Insulin receptor	R_{1145} D I Y* E T D Y* Y* R K G G K G	Autophosphorylation (mouse)	61, 64
	R_{1143} D I Y* E T D Y* Y* R K G G K G	Autophosphorylation (human)	64, 113
	K_{1313} R S Y* E E H I P Y* T H M N G G K	Autophosphorylation (human)	64, 113
	E N F D D Y* M K E	pp15 [422(aP2) protein]	114
	D_{93} K D G N G Y* I S A A E	Calmodulin	67, 68
PDGF receptor	D_{746} E S V D Y* V P M L D M K	Autophosphorylation	65
	D_{854} S N Y* I S K	Autophosphorylation	65
pp50$^{v\text{-}abl}$	I_{14} E N E E Q E Y* V Q T V K S S	Lipocortin I (p35, calpactin II), Human	81
pp60$^{c\text{-}src}$	L_{308} E E E E E E Y* M P M E D L Y	Polyomavirus middle T antigen	61, 93
	S_{243} L L S N P T Y* S V M R S H S	Polyomavirus middle T antigen[#]	61
	I_{14} E N E E Q E Y* V Q T V K S S	Lipocortin I (p35, calpactin II), human	81
	R_{409} L I E D N E Y* T A R Q G A K	Autophosphorylation	143

(*continued*)

TABLE I (*continued*)

Protein kinase	Phosphorylation site sequence	Protein	Refs.[b]
pp60$^{c\text{-}src}$ and/or unknown kinase(s)	F_{520} T S T E P Q Y* Q P G E N L	pp60$^{c\text{-}src}$	61, 66, 143
pp60$^{v\text{-}src}$	R_{409} L I E D N E Y* T A R Q G A K	Autophosphorylation	61
	S_{336} G G K G G S Y* S Q A A C S D	HLA-B7 (α chain)#	61
	S_{290} D R K G G S Y* S Q A A S S D	HLA-A2 (α chain)#	61
	H_{16} S T P P S A Y* G S V K A Y T	Lipocortin II (p36, calpactin I heavy chain)	61, 62
p90$^{gag\text{-}yes}$	R_{417} L I E D N E Y* T A R Q G A K	Autophosphorylation# (residue numbers based on human sequence)	61
p56^{lck}	R_{387} L I E D N E Y* T A R E G A K	Autophosphorylation	61
p140$^{gag\text{-}fps}$	R_{417} Q E E D G V Y* A S T G G M K	Autophosphorylation	61
	K_{231} Q V V E S A Y* E V I R L K G	Lactate dehydrogenase# (residue numbers based on chicken sequence)	61
	S_{36} G A S T G I Y* E A L E L R	Enolase# (residue numbers based on human sequence)	61
p110$^{gag\text{-}fes}$	R_{493} E A A D G I Y* A A S G G L R	Autophosphorylation	61, 149
p85$^{gag\text{-}fes}$	R_{361} E E A D G V Y* A A S G G L R	Autophosphorylation	61, 149
p120$^{gag\text{-}abl}$	R_{386} L M T G D T Y* T A H A G A G	Autophosphorylation# (residue numbers based on human sequence)	61
Endogenous kinase (?p40)	M_1 E E L Q D D Y* E D D M E E N	Band 3	61, 74, 75
p40	E_{669} E D G E R Y* D E D E E	Glycogen synthase (residue numbers based on human muscle sequence)	74, 150
Unknown kinase(s)	F_{498} T A T E G Q Y* Q P Q P	p56^{lck}	143
Unknown kinase	K_{36} K R K S* G N S R E R	Avian retrovirus nucleocapsid protein (pp12)	77
Unknown kinase	K_9 I G E G T Y* G V V Y K A R H K	cdc2$^+$ (pp34)	106

[a] #, Residue numbers obtained from SWISSPROT protein data base; (P) denotes a phosphorylated residue which acts as a substrate specificity determinant.

[3]

Refs.[b]

66, 143

62

149

149

74, 75

150

orylated

[b] References for Tables I and II are combined at the end of the chapter.
[c] The phosphorylation sites for the β-adrenergic receptor kinase are inferred from mutagenesis studies.[50]
[d] Artificial protein substrates have not been included for casein kinase II.
[e] Only selected phosphorylation sites for cAMP-dependent protein kinase and protein kinase C have been included to illustrate the various recognition motifs.
[f] P. J. Robinson (unpublished result, 1990).
[g] Although the growth-associated H1 histone kinases appear to require proline residues for substrate recognition, so do a number of other kinases, including glycogen synthase kinase-3, growth factor-regulated kinase, mitogen-activated protein kinase, proline-directed protein kinase, and sperm-specific histone kinase.
[h] Threonine residues 340 and 342 are assumed sites based on stoichiometry of rhodopsin phosphorylation.
[i] Some doubt exists as to whether tyrosine phosphorylation of the various tyrosine kinases is due to autophosphorylation *in vivo*.
[j] Residue numbers for protein phosphatase-1 G subunit obtained from P. Tang, J. Bondor, and A. A. DePaoli-Roach (personal communication, 1990).

TABLE II
CONSENSUS PHOSPHORYLATION SITES: SPECIFICITY MOTIFS FOR PROTEIN KINASES

Protein kinase	S:T ratio[a]	Motif[b]	Frequency
Calmodulin-dependent protein kinase II	15:5	XRXXS*/T*	13/20[c]
		XRXXS*/T*V	6/20
Casein kinase I	8:1	S(P)XXS*/T*	5/9[d]
Casein kinase II	28:2	S*/T*XXEX	23/30
		S*/T*XXDX	3/20[e]
cAMP-dependent protein kinase[f]	40:6	RXS*	21/46
		RRXS*	12/46
		RXXS*	11/46
		KRXXS*	2/46
cGMP-dependent protein kinase	7:3[g]	R/KXS*/T*	9/10
		R/KXXS*/T*	8/10
		R/KR/KXS*/T*	7/10
		R/KXXXS*/T*	5/10
		S*/T*XR/K	2/10
Glycogen synthase kinase-3	10:2	S*XXXS(P)	6/12[h]
Growth-associated histone H1 kinase (MPF, cdc2+/CDC28 protein kinases)	7:8	S*/T*PXK/R	6/15
		K/RS*/T*P	5/15
		S*/T*PK/R	4/15
Phosphorylase kinase	3:0	K/RXXS*V/I	3/3
Protein kinase C	31:6	S*/T*XK/R	20/37[i]
		K/RXXS*/T*	13/37
		K/RXXS*/T*XK/R	7/37
		K/RXS*/T*	10/37
		K/RXS*/T*XK/R	6/37

(continued)

TABLE II (*continued*)

Protein kinase	S:T ratio[a]	Motif[b]	Frequency
Tyrosine kinase[j]			
EGF-receptor kinase	Tyrosine	XE/DY*X	7/14
		XE/DY*I/L/V	5/14

[a] S:T ratio is for the total number of phosphorylation sites.
[b] Asterisks indicate the phosphorylated residue. Specificity determinants are shown in bold type.
[c] Three of 20 phosphorylation sites for calmodulin-dependent protein kinase II are on threonine residues, including both autophosphorylation sites.
[d] Assuming phosphorylation of Ser-3 in glycogen synthase directs Ser-7 phosphorylation.
[e] Eighteen of 30 Casein kinase II phosphorylation site sequences contain 3 consecutive acidic residues following the phosphorylated residue.
[f] cAMP-dependent protein kinase motifs and frequency data are derived from O. Zetterqvist, U. Ragnarsson, and L. Engström, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 171. Uniscience CRC Press, Boca Raton, Florida, 1990, and from Table I in this chapter. The most striking feature of the cAMP-dependent protein kinase phosphorylation site sequences is the variability, with less than one-third corresponding to the **RRXS*** motif.
[g] The only autophosphorylation site included was Thr-58. Autophosphorylation at Ser-50, Ser-72,and Thr-84 occurs only following activation with cAMP.
[h] Assuming sequential phosphorylation of glycogen synthase and protein phosphatase-1 G subunit by glycogen synthase kinase-3.
[i] Twenty-three of 37 protein kinase C phosphorylation site sequences contain an adjacent hydrophobic residue on the COOH-terminal side of the phosphorylated residue.
[j] Apart from the EGF-receptor kinase, the tyrosine kinase substrate sequences do not reveal consensus recognition motifs. Relatively few phosphorylation site sequences are known for exogenous proteins and autophosphorylation sites may not reflect the specificity determinants required in substrates. A more informative indication of likely phosphorylation site arrangements may be drawn from this volume [10].

References

[1] O. Zetterqvist, U. Ragnarsson, and L. Engström, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 171. Uniscience CRC Press, Boca Raton, Florida, 1990.
[2] J. R. Cherry, T. R. Johnson, C. Dollard, J. R. Shuster, and C. L. Denis, *Cell (Cambridge, Mass.)* **56**, 409 (1989).
[3] D. B. Glass, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 209. Uniscience CRC Press, Boca Raton, Florida, 1990.
[4] K.-F. J. Chan, M. O. Hurst, and D. J. Graves, *J. Biol. Chem.* **257**, 3655 (1982).
[5] R. B. Pearson, R. Jakes, M. John, and J. Kendrick-Jones, *FEBS Lett.* **168**, 108 (1984).
[6] G. Matsuda, Y. Suzuyama, T. Maita, and T. Umegane, *FEBS Lett.* **84**, 53 (1977).
[7] A. Kishimoto, K. Nishiyama, H. Nakanishi, Y. Uratsuji, H. Nomura, Y. Takeyama, and Y. Nishizuka, *J. Biol. Chem.* **260**, 12492 (1985).
[8] O. Koshio, Y. Akanuma, and M. Kasuga, *FEBS Lett.* **254**, 22 (1989).

[9] R. J. Davis, G. L. Johnson, D. J. Kelleher, J. K. Anderson, J. E. Mole, and M. P.Czech, *J. Biol. Chem.* **261,** 9034 (1986).
[10] M. Ikebe, D. J. Hartshorne, and M. Elzinga, *J. Biol. Chem.* **262,** 9569 (1987).
[11] J. M. Graff, D. J. Stumpo, and P. J. Blackshear, *J. Biol. Chem.* **264,** 11912 (1989).
[12] S. Jakes, T. G. Hastings, E. M. Reimann, and K. K. Schlender, *FEBS Lett.* **234,** 31 (1988).
[13] Reference deleted in proof.
[14] A. J. Czernik, D. T. Pang, and P. Greengard, *Proc. Natl. Acad. Sci. U.S.A.* **84,** 7518 (1987).
[15] H. K. B. Simmerman, J. H. Collins, J. L. Theibert, A. D. Wegener, and L. R. Jones, *J. Biol. Chem.* **261,** 13333 (1986).
[16] J. R. Woodgett, M. T. Davison, and P. Cohen, *Eur. J. Biochem.* **136,** 481 (1983).
[17] Y. Hashimoto and T. R. Soderling, *J. Biol. Chem.* **264,** 16524 (1989).
[18] D. J. Campbell, D. G. Hardie, and P. R. Vulliet, *J. Biol. Chem.* **261,** 10489 (1986).
[19] A. P. Doskeland, C. M. Schworer, S. O. Doskeland, T. D. Chrisman, T. R. Soderling, J. R. Corbin, and T. Flatmark, *Eur. J. Biochem.* **145,** 31 (1984).
[20] T. R. Soderling, C. M. Schworer, M. R. El-Maghrabi, and S. J. Pilkis, *Biochem. Biophys. Res. Commun.* **139,** 1017 (1986).
[21] S. Shoji, J. Ohnishi, T. Funakoshi, K. Fukunaga, E. Miyamoto, H. Ueki, and Y. Kubota, *J. Biochem. (Tokyo)* **102,** 1113 (1987).
[22] A. C. Nairn and H. C. Palfrey, *J. Biol. Chem.* **262,** 17299 (1987).
[23] T. Hunter, N. Ling, and J. A. Cooper, *Nature (London)* **311,** 480 (1984).
[24] C. House, R. E. H. Wettenhall, and B. E. Kemp, *J. Biol. Chem.* **262,** 772 (1987).
[25] J.-A. Girault, H. C. Hemmings, K. R. Williams, A. C. Nairn, and P. Greengard, *J. Biol. Chem.* **264,** 21748 (1989).
[26] S. P. Lees-Miller and C. W. Anderson, *J. Biol. Chem.* **264,** 2431 (1989).
[27] Y. Tashiro, S. Matsumura, N. Murakami, and A. Kumon, *Arch. Biochem. Biophys.* **233,** 540 (1984).
[28] T. A. J. Haystead, D. G. Campbell, and D. G. Hardie, *Eur. J. Biochem.* **175,** 347 (1988).
[29] L. A. Pinna, F. Meggio, and F. Marchiori, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 145. Uniscience CRC Press, Boca Raton, Florida, 1990.
[30] H. Flotow and P. J. Roach, *J. Biol. Chem.* **264,** 9126 (1989).
[31] E. Erikson and J. L. Maller, *Second Messengers Phosphoproteins* **12,** 135 (1988).
[32] H. Brzeska, T. J. Lynch, B. Martin, and E. D. Korn, *J. Biol. Chem.* **264,** 19340 (1989).
[33] P. R. Vulliet, F. L. Hall, J. P. Mitchell, and D. G. Hardie, *J. Biol. Chem.* **264,** 16292 (1989).
[34] J. L. Countaway, I. C. Northwood, and R. J. Davis, *J. Biol. Chem.* **264,** 10828 (1989).
[35] D. B. Rylatt, A. Aitken, T. Bilham, G. D. Condon, N. Embi, and P. Cohen, *Eur. J. Biochem.* **107,** 529 (1980).
[36] P. Dent, D. G. Campbell, M. J. Hubbard, and P. Cohen, *FEBS Lett.* **248,** 67 (1989).
[37] S. P. Lees-Miller and C. W. Anderson, *J. Biol. Chem.* **264,** 17275 (1989).
[38] D. R. Colthurst, D. G. Campbell, and C. G. Proud, *Eur. J. Biochem.* **166,** 357 (1987).
[39] W. Kudlicki, R. E. H. Wettenhall, B. E. Kemp, R. Szyszka, G. Kramer, and B. Hardesty, *FEBS Lett.* **215,** 16 (1987).
[40] W. Rychlik, M. A. Russ, and R. E. Rhoads, *J. Biol. Chem.* **262,** 10434 (1987).
[41] P. Cohen, D. Yellowlees, A. Aitken, A. Donella-Deana, B. A. Hemmings, and P. J. Parker, *Eur. J. Biochem.* **124,** 21 (1982).
[42] D. G. Hardie, D. Carling, and A. T. R. Sim, *Trends Biochem. Sci.* **14,** 20 (1989).
[43] D. G. Hardie, D. Carling, S. Ferrari, P. S. Guy, and A. Aitken, *Eur. J. Biochem.* **157,** 553 (1986).

[44] P. Thompson and J. B. C. Findlay, *Biochem. J.* **220**, 773 (1984).
[45] R. Paxton, M. Kuntz, and R. A. Harris, *Arch. Biochem. Biophys.* **244**, 187 (1986).
[46] K. Montgomery and A. S. Mak, *J. Biol. Chem.* **259**, 5555 (1984).
[47] A. C. Borthwick, W. H. Holmes, and H. G. Nimmo, *FEBS Lett.* **174**, 112 (1984).
[48] R. A. Masaracchia, B. E. Kemp, and D. A. Walsh, *J. Biol. Chem.* **252**, 7109 (1977).
[49] P. R. Clarke and D. G. Hardie, *EMBO J.* **9**, 2439 (1990).
[50] W. P. Hausdorff, M. Bouvier, B. F. O'Dowd, G. P. Irons, M. G. Caron, and R. J. Lefkowitz, *J. Biol. Chem.* **264**, 12657 (1989).
[51] A. M. Edelman, D. K. Blumenthal, and E. G. Krebs, *Annu. Rev. Biochem.* **56**, 567 (1987).
[52] N. A. Morrice and R. E. H. Wettenhall, personal communication (1990).
[53] T. A. Langan, *Methods Cell Biol.* **19**, 127 (1978).
[54] T. A. Langan, J. Gautier, M. Lohka, R. Hollingsworth, S. Moreno, P. Nurse, J. Maller, and R. A. Sclafani, *Mol. Cell. Biol.* **9**, 3860 (1989).
[55] H. C. Hemmings, J.-A. Girault, K. R. Williams, M. B. LoPresti, and P. Greengard, *J. Biol. Chem.* **264**, 7726 (1989).
[56] N. Murakami, G. Healy-Louie, and M. Elzinga, *J. Biol. Chem.* **265**, 1041 (1990).
[57] C. S. Hill, L. C. Packman, and J. O. Thomas, *EMBO J.* **9**, 805 (1990).
[58] C. E. M. Voorter, J. W. M. Mulders, H. Bloemendal, and W. W. de Jong, *Eur. J. Biochem.* **160**, 203 (1986).
[59] C.E. M. Voorter, W. A. de Haard-Hoekman, E. S. Roersma, H. E. Meyer, H. Bloemendal, and W. W. de Jong, *FEBS Lett.* **259**, 50 (1989).
[60] M. Ikebe, D. J. Hartshorne, and M. Elzinga, *J. Biol. Chem.* **261**, 36 (1986).
[61] R. L. Geahlen and M. L. Harrison, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 239. Uniscience CRC Press, Boca Raton, Florida, 1990.
[62] J. R. Glenney and B. F. Tack, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7884 (1985).
[63] B. K. De, K. S. Misono, T. J. Lukas, B. Mroczkowski, and S. Cohen, *J. Biol. Chem.* **261**, 13784 (1986).
[64] J. R. Flores-Riveros, E. Sibley, T. Kastelic, and M. D. Lane, *J. Biol. Chem.* **264**, 21557 (1989).
[65] A. Kazlauskas and J. A. Cooper, *Cell (Cambridge, Mass.)* **58**, 1121 (1989).
[66] M. Okada and H. Nakagawa, *J. Biol. Chem.* **264**, 20886 (1989).
[67] J. P. Laurino, J. C. Colca, J. D. Pearson, D. B. DeWald, and J. M. McDonald, *Arch. Biochem. Biophys.* **265**, 8 (1988).
[68] D. M. Watterson, F. Sharief, and T. C. Vanaman, *J. Biol. Chem.* **255**, 962 (1980).
[69] B. E. Kemp, *J. Biol. Chem.* **254**, 2638 (1979).
[70] S. J. Yeaman, *Biochim. Biophys. Acta* **1052**, 128 (1990).
[71] C. F. B. Holmes, J. Kuret, A. A. K. Chisholm, and P. Cohen, *Biochim. Biophys. Acta* **870**, 408 (1986).
[72] B. A. Hemmings, A. Aitken, P. Cohen, M. Rymond, and F. Hofmann, *Eur. J. Biochem.* **127**, 473 (1982).
[73] S. M. Van Patten, G. J. Heisermann, H.-C. Cheng, and D. A. Walsh, *J. Biol. Chem.* **262**, 3398 (1987).
[74] A. M. Mahrenholz, P. Votaw, P. J. Roach, A. A. Depaoli-Roach, T. F. Zioncheck, M. L. Harrison, and R. L. Geahlen, *Biochem. Biophys. Res. Commun.* **155**, 52 (1988).
[75] S. A. Dekowski, A. Rybicki, and K. Drickamer, *J. Biol. Chem.* **258**, 2750 (1983).
[76] D. B. Bylund and E. G. Krebs, *J. Biol. Chem.* **250**, 6355 (1975).
[77] X. Fu, P. T. Tuazon, J. A. Traugh, and J. Leis, *J. Biol. Chem.* **263**, 2134 (1988).
[78] J. Fujii, A. Ueno, K. Kitano, S. Tanaka, M. Kadoma, and M. Tada, *J. Clin. Invest.* **79**, 301 (1987).
[79] T. Patschinsky, T. Hunter, and B. M. Sefton, *J. Virol.* **591**, 73 (1986).

[80] C. H. Michnoff, B. E. Kemp, and J. T. Stull, *J. Biol. Chem.* **261**, 8320 (1986).
[81] L. Varticovski, S. B. Chahwala, M. Whitman, L. Cantley, D. Schindler, E. P. Chow, L. K. Sinclair, and R. B. Pepinsky, *Biochemistry* **27**, 3682 (1988).
[82] D. D. Schlaepfer and H. T. Haigler, *Biochemistry* **27**, 4253 (1988).
[83] K. L. Gould, J. R. Woodgett, C. M. Isacke, and T. Hunter, *Mol. Cell. Biol.* **6**, 2738 (1986).
[84] P. Heldin and E. Humble, *Arch. Biochem. Biophys.* **252**, 49 (1987).
[85] A. Safran, R. Sagi-Eisenberg, D. Neumann, and S. Fuchs, *J. Biol. Chem.* **262**, 10506 (1987).
[86] S. Kitamura, S. Ando, M. Shibata, K. Tanabe, C. Sato, and M. Inagaki, *J. Biol. Chem.* **264**, 5674 (1989).
[87] B. Luscher, E. Christensen, D. W. Litchfield, E. G. Krebs, and R. N. Eisenman, *Nature (London)* **344**, 517 (1990).
[88] B. Luscher, E. A. Kuenzel, E. G. Krebs, and R. N. Eisenman, *EMBO J.* **8**, 1111 (1989).
[89] P. Cohen, *Nature (London)* **296**, 613 (1982).
[90] H. Olsson, P. Stralfors, and P. Belfrage, *FEBS Lett.* **209**, 175 (1986).
[91] A. J. Garton, D. G. Campbell, D. Carling, D. G. Hardie, R. J. Colbran, and S. J. Yeaman, *Eur. J. Biochem.* **179**, 249 (1989).
[92] J. W. Kim, S. S. Sim, U.-H. Kim, S. Nishibe, M. I. Wahl, G. Carpenter, and S. G. Rhee, *J. Biol. Chem.* **265**, 3940 (1990).
[93] D. A. Talmage, R. Freund, A. T. Young, J. Dahl, C. J. Dawe, and T. L. Benjamin, *Cell (Cambridge, Mass.)* **59**, 55 (1989).
[94] B. L. Margolis, I. Lax, R. Kris, M. Dombalagian, A. M. Honegger, R. Howk, D. Givol, A. Ullrich, and J. Schlessinger, *J. Biol. Chem.* **264**, 10667 (1989).
[95] S. Shoji, D. C. Parmelee, R. D. Wade, S. Kumar, L. H. Ericsson, K. A. Walsh, H. Neurath, G. L. Long, J. G. Demaille, E. H. Fischer, and K. Titani, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 848 (1981).
[96] C. R. Lin, M. S. Kapiloff, S. Durgerian, K. Tatemoto, A. F. Russo, P. Hanson, H. Schulman, and M. G. Rosenfeld, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5962 (1987).
[97] G. Thiel, A. J. Czernik, F. Gorelick, A. C. Nairn, and P. Greengard, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6337 (1987).
[98] C. M. Schworer, R. J. Colbran, J. R. Keefer, and T. R. Soderling, *J. Biol. Chem.* **263**, 13486 (1988).
[99] M. K. Bennett and M. B. Kennedy, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1794 (1987).
[100] F. L. Hall, J. P. Mitchell, and P. R. Vulliet, *J. Biol. Chem.* **265**, 6944 (1990).
[101] M. M. King, T. J. Fitzgerald, and G. M. Carlson, *J. Biol. Chem.* **258**, 9925 (1983).
[102] T. J. Singh, A. Akutsuka, and K.-P. Huang, *J. Biol. Chem.* **257**, 13379 (1982).
[103] C. J. Fiol, A. Wang, R. W. Roeske, and P. J. Roach, *J. Biol. Chem.* **265**, 6061 (1990).
[104] S. Shenoy, J.-K. Choi, S. Bagrodia, T. D. Copeland, J. L. Maller, and D. Shalloway, *Cell (Cambridge, Mass.)* **57**, 763 (1989).
[105] L. J. Cisek and J. L. Corden, *Nature (London)* **339**, 679 (1989).
[106] K. L. Gould and P. Nurse, *Nature (London)* **342**, 39 (1989).
[107] A. K. Erickson, D. M. Payne, P. A. Martino, A. J. Rossomando, J. Shabanowitz, M. J. Weber, D. F. Hunt, and T. W. Sturgill, *J. Biol. Chem.* **265**, 19728 (1990).
[108] D. F. Carmichael, R. L. Geahlen, S. M. Allen, and E. G. Krebs, *J. Biol. Chem.* **257**, 10440 (1982).
[109] B. L. Hill, K. Drickamer, F. M. Brodsky, and P. Parham, *J. Biol. Chem.* **263**, 5499 (1988).
[110] G. M. C. Janssen, G. D. F. Maessen, R. Amons, and W. Moller, *J. Biol. Chem.* **263**, 11063 (1988).
[111] G. M. Walton, W. S. Chen, M. G. Rosenfeld, and G. N. Gill, *J. Biol. Chem.* **265**, 1750 (1990).

[112] Reference deleted in proof.
[113] H. E. Tornqvist, J. R. Gunsalus, R. A. Nemenoff, A. R. Frackelton, M. W. Pierce, and J. Avruch, *J. Biol. Chem.* **263**, 350 (1988).
[114] R. C. Hresko, M. Bernier, R. D. Hoffman, J. R. Flores-Riveros, K. Liao, D. M. Laird, and M. D. Lane, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8835 (1988).
[115] R. E. Lewis, G. P. Wu, R. G. MacDonald, and M. P. Czech, *J. Biol. Chem.* **265**, 947 (1990).
[116] S. Nakajo, Y. Masuda, K. Nakaya, and Y. Nakamura, *J. Biochem. (Tokyo)* **104**, 946 (1988).
[117] M. S. Barbosa, C. Edmonds, C. Fisher, J. T. Schiller, D. R. Lowy, and K. H. Vousden, *EMBO J.* **9**, 153 (1990).
[118] J. Kuret, J. R. Woodgett, and P. Cohen, *Eur. J. Biochem.* **151**, 39 (1985).
[119] T. Hunter, P. Angel, W. J. Boyle, R. Chiu, E. Freed, K. L. Gould, C. M. Isacke, M. Karin, R. A. Lindberg, and P. van der Geer, *Cold Spring Harbor Symp. Quant. Biol.* **53**, 131 (1988).
[120] W. J. Boyle, T. Smeal, L. H. K. Defize, P. Angel, J. R. Woodgett, M. Karin, and T. Hunter, *Cell (Cambridge,Mass.)* **64**, 573 (1991).
[121] R. E. Lewis, L. Cao, D. Perregaux, and M. P. Czech, *Biochemistry* **29**, 1807 (1990).
[122] A. Ullrich, J. R. Bell, E. Y. Chen, R. Herrera, R. M. Petruzzelli, T. J. Dull, A. Gray, L. Coussens, Y.-C. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfeld, O. M. Rosen, and J. Ramachandran, *Nature (London)* **313**, 756 (1985).
[123] J. R. Woodgett, K. L. Gould, and T. Hunter, *Eur. J. Biochem.* **161**, 177 (1986).
[124] D. A. Shackelford and I. S. Trowbridge, *J. Biol. Chem.* **261**, 8334 (1986).
[125] K. L. Gould, J. R. Woodgett, J. A. Cooper, J. E. Buss, D. Shalloway, and T. Hunter, *Cell (Cambridge, Mass.)* **42**, 849 (1985).
[126] S. J. Decker, B. Dorai, and S. Russell, *J. Virol.* **62**, 3649 (1988).
[127] J. Downward, Y. Yarden, E. Mayes, G. Scrace, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger, and M. D. Waterfield, *Nature (London)* **307**, 521 (1984).
[128] J. Rittenhouse, L. Moberly, and F. Marcus, *J. Biol. Chem.* **262**, 10114 (1987).
[129] J. Kuret, K. E. Johnson, C. Nicolette, and M. J. Zoller, *J. Biol. Chem.* **263**, 9149 (1988).
[130] Y. Hashimoto and T. R. Soderling, *Arch. Biochem. Biophys.* **278**, 41 (1990).
[131] N. J. Olson, R. B. Pearson, D. S. Needleman, M. Y. Hurwitz, B. E. Kemp, and A. R. Means, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2284 (1990).
[132] P. Tapley, A. Kazlauskas, J. A. Cooper, and L. R. Rohrschneider, *Mol. Cell. Biol.* **10**, 2528 (1990).
[133] P. van der Geer and T. Hunter, *Mol. Cell. Biol.* **10**, 2991 (1990).
[134] Y. Goldberg, C. Glineur, J.-C. Gasquiere, A. Ricouart, J. Sap, B. Vennstrom, and J. Ghysdael, *EMBO J.* **7**, 2425 (1988).
[135] Z. Ahmad, F.-T. Lee, A. DePaoli-Roach, and P. J. Roach, *J. Biol. Chem.* **259**, 8743 (1984).
[136] E. D. Apel, M. F. Byford, D. Au, K. A. Walsh, and D. R. Storm, *Biochemistry* **29**, 2330 (1990).
[137] J. Pines and T. Hunter, *New Biol.* **2**, 389 (1990).
[138] E. T. Kipreos and J. Y. J. Wang, *Science* **248**, 217 (1990).
[139] T. M. Martensen, B. M. Martin, and R. L. Kincaid, *Biochemistry* **28**, 9243 (1989).
[140] G. M. Walton, J. Spiess, and G. N. Gill, *J. Biol. Chem.* **260**, 4745 (1985).
[141] T. A. J. Haystead and D. G. Hardie, *Eur. J. Biochem.* **175**, 339 (1988).
[142] A. Ullrich, L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Wittle, M. D. Waterfield, and P. H. Seeburg, *Nature (London)* **309**, 418 (1984).

[143] J. A. Cooper, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 85. Uniscience CRC Press, Boca Raton, Florida, 1990.
[144] H. U. Beuscher, M. W. Nickells, and H. R. Colten, *J. Biol. Chem.* **263**, 4023 (1988).
[145] G. J. Heisermann and G. N. Gill, *J. Biol. Chem.* **263**, 13152 (1988).
[146] M. L. Tremblay, C. J. McGlade, G. E. Gerber, and P. E. Branton, *J. Biol. Chem.* **263**, 6375 (1988).
[147] P. Jeno, L. M. Ballou, I. Novak-Hofer, and G. Thomas, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 406 (1988).
[148] J. Krieg, J. Hofsteenge, and G. Thomas, *J. Biol. Chem.* **263**, 11473 (1988).
[149] A. Hampe, I. Laprevotte, F. Galibert, L. A. Fedele, and C. J. Sherr, *Cell (Cambridge, Mass.)* **30**, 775 (1982).
[150] M. F. Browner, K. Nakano, A. G. Bang, and R. J. Fletterick, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1443 (1989).

EXHIBIT 2

[10] Design and Use of Peptide Substrates for Protein Kinases

By BRUCE E. KEMP and RICHARD B. PEARSON

Synthetic peptide substrates have played an important role in the study of protein kinase (PK) substrate specificity as well as in the measurement of protein kinase activities in cell extracts. Their great utility as experimental reagents became apparent from studies of the substrate specificity of the cAMP-dependent protein kinases and phosphorylase kinase.[1] It was found that the principal substrate specificity determinants for these enzymes were located in short segments of the primary sequence around phosphorylation sites. Arginine residues were identified as important specificity determinants for both of these enzymes by studies using genetic variants of protein substrates[1] and synthetic peptides.[2] Some short synthetic peptides of 7–10 residues were phosphorylated with kinetic properties comparable to the parent protein. The sequencing of a number of phosphorylation sites from a variety of protein substrates for the cAMP-dependent protein kinase provided evidence for the concept of a recognition motif, RRXS or KRXXS, indicating that all of the features necessary for phosphorylation could be combined in a short peptide.[2a] While these findings certainly reinforced the concept of a recognition motif it should be noted that only approximately one-third of the known phosphorylation sites for the cAMP-dependent protein kinase actually conform to the RRXS or KRXXS motif.[3] The liver pyruvate kinase peptide LRRASLG[3] (kemptide) proved to be an excellent substrate for the cAMP-dependent protein kinase with a K_m of approximately 5 μM and a V_{max} of 16 μmol min^{-1} mg^{-1}. Subsequently it was found that synthetic peptides corresponding to other phosphorylation site sequences for this enzyme varied widely in their capacity to act as substrates with K_m values in the range 0.1 to 2 mM. At present this phenomenon is understood in terms of the local phosphorylation site

[1] B. E. Kemp, D. B. Bylund, T. S. Huang, and E. G. Krebs, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3448 (1975).

[2] B. E. Kemp, D. J. Graves, E. Benjamini, and E. G. Krebs, *J. Biol. Chem.* **252**, 4888 (1977).

[2a] Single-letter abbreviations are used for amino acids: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; X, unknown; Y, tyrosine.

[3] O. Zetterqvist, U. Ragnarsson, and L. Engström, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 171. Uniscience CRC Press, Boca Raton, Florida, 1990.

sequence being the principal determinant of substrate specificity, but its context in the parent protein structure is also important. It seems likely that nearby residues required to accommodate the phosphorylation site sequence into the parent protein structure may influence the capacity of the corresponding synthetic peptide to act as a substrate either positively or negatively.

The major goals of designing peptide substrates for protein kinases are to construct peptides that have excellent kinetic properties and a high degree of specificity. Although some synthetic peptides are phosphorylated with K_m and V_{max} values comparable to their parent proteins, many are not and the reason for this is poorly understood. Peptide substrates that are phosphorylated with K_m values of less than 50 μM and preferably in the range 1 to 10 μM are typically the most useful. A low K_m value improves the likelihood that the peptide substrate will be relatively specific for that enzyme and has a cost benefit if large numbers of assays are required using an expensive synthetic substrate. On the other hand, from a practical viewpoint the V_{max} value is more important than the K_m value because it determines the detection sensitivity of the peptide phosphorylation reaction. The greater the V_{max} value, the greater the latitude available for varying the peptide substrate concentration. Overlapping specificities of some protein kinases occurs, such as for phosphorylase kinase, protein kinase C, and the multifunctional calmodulin-dependent protein kinase, all of which phosphorylate Ser-7 in the glycogen synthase peptide (see below). Examples will be given below of ways in which the sensitivity and specificity of protein kinase peptide substrates have been enhanced. Sensitivity is a function of V_{max}/K_m and the relative specificity for a peptide substrate toward two enzymes is a function of the ratio, [V_{max}/K_m (enzyme 1)]/[V_{max}/K_m (enzyme 2)], or coefficient of specificity.

Peptide Substrate Synthesis

Previously Glass[4] has reported procedures for the synthesis of oligopeptides for the study of cyclic nucleotide-dependent protein kinases in an earlier volume of this series. For this reason a detailed account of the chemistry of oligopeptide synthesis will not be covered here. The Merrifield solid-phase synthesis procedure with either tBoc or Fmoc chemistries is used widely.[5] For the chemical synthesis of phosphorylated peptides, see [18] and [19] (Volume 201 in this series). The advent of modern automated peptide synthesizers such as the model 430 (Applied Biosystems, Foster

[4] D. B. Glass, this series, V l. 99, p. 119.
[5] S. B. H. Kent, *Annu. Rev. Biochem.* 57, 957 (1988).

City, CA) has meant that the assembly of peptides is no longer rate limiting. However, peptide purification and characterization remain areas of great importance. The assembled peptides are typically cleaved from the resin using anhydrous HF[6] or trifluoromethanesulfonic acid. The resultant crude peptide is then extracted from the resin using one of several suitable volatile buffers, depending on the amino acid composition of the peptide. Most basic and neutral peptides are soluble in 60% acetonitrile (v/v) containing 0.1% trifluoroacetic acid (v/v) and acidic peptides are soluble in 60% acetonitrile (v/v) containing 0.1 *M* NH_4HCO_3. Either acetic acid (30%, v/v) or trifluoroacetic acid (50%, v/v) can also be used for more hydrophobic peptides. The completeness of peptide extraction from the resin is readily monitored using the resin ninhydrin test.[7]

In general, protein kinase substrates are relatively polar and solubility is not normally limiting. The most reliable way to purify synthetic peptide substrates routinely is to use both ion-exchange and reversed-phase chromatography. Alternatively, a two-step reversed-phase chromatography using ion pairing at two pH values has been widely used for other synthetic peptides.[8] Reversed-phase chromatography carried out at a single pH can be insufficient, particularly in cases where hydrophobic and hydrophilic residues are separated to different ends of the peptide. Since many protein kinases utilize either acidic or basic residues as specificity determinants an ion-exchange step in the purification of peptide substrates is particularly effective. The procedures outlined below are designed for a single person simultaneously synthesizing and purifying peptides in a small laboratory with limited instrument resources.

Ion-exchange chromatography is conveniently carried out using a commercial purification system such as fast protein liquid chromatography (Pharmacia Piscataway, NJ FPLC) or Waters (Milford, MA) 650 protein purification system (K. I. Mitchelhill and B. E. Kemp, unpublished). Basic peptides are purified on a S-Sepharose HP (Pharmacia) cation-exchange column (1 × 25 cm) using a 0 to 1.0 *M* NaCl gradient in trifluoroacetic acid 0.1% (v/v) with 20% (v/v) acetonitrile. The inclusion of acetonitrile in the ion-exchange chromatography buffer is to enhance peptide solubility and recovery. The flow rate is 1 ml/min for 1140 min (overnight). The peptide peak is located using ninhydrin staining of dot blots of the fractions followed by analytical reversed-phase high-performance liquid chromatography (HPLC) of the fractions containing peptide. The analytical separa-

[6] J. M. Stewart and J. D. Young, "Solid Phase Peptide Synthesis," pp. 44 and 66. Freeman, San Francisco, California, 1966.

[7] V. K. Savin, S. B. H. Kent, J. P. Tam, and R. B. Merrifield, *Anal. Biochem.* **117**, 147 (1981).

[8] J. Rivier, *J. Liq. Chromatogr.* **1**, 343 (1978).

tions are routinely carried out on an RP-300 guard column (4.6 × 30 mm, 7-μm particle size) using a linear gradient from 0 to 60% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid. This gradient may be modified to optimize the elution time and resolution, facilitating a larger throughput of samples. The peptide is detected by monitoring absorbance at 210 nm. The fractions are pooled and the acetonitrile removed by rotary evaporation. The ion-exchange purified peptide is chromatographed on a hand-packed preparative reversed-phase low-pressure column (2.5 × 100 cm) containing Vydac C_{18} resin (218TPB 2030). The combination of a large column and overnight chromatography provides resolution equivalent to an analytical HPLC column. An acetonitrile gradient in trifluoroacetic acid (0.1%, v/v) is run at 1.5 ml/min for 14 hr. The acetonitrile gradient is tailored to the properties of the particular peptide based on its retention on analytical reversed-phase HPLC. Again the fractions containing peptide are located by dot blot developed with ninhydrin followed by analytical reversed-phase HPLC of the fractions containing peptide. On-line UV monitoring of preparative column eluants is an advantage, but in many cases the concentration of eluted peptide may be off scale on the detector unless customized preparative flow cells are used. The fractions containing pure peptide are pooled and lyophilized as their trifluoroacetic acid salts.

Acidic peptides are purified on a Q-Sepharose HP (Pharmacia) anion-exchange column (1.6 × 10 cm) using a 0 to 0.4 *M* NaCl gradient in 20 m*M* NH_4HCO_3 with 20% acetonitrile (v/v) and a flow rate of 1 ml/min for 1140 min. The acetonitrile is removed by rotary evaporation prior to reversed-phase chromatography of the acidic peptides in the presence of 20 m*M* NH_4HCO_3. The anion-exchange column fractions can be screened for peptide using the ninhydrin dot blot procedure provided the blot is dipped in diisopropylethylamine (30%, v/v, in dichloromethane) and dried to removed NH_4^+ prior to staining with ninhydrin. By using a combination of ion-exchange and reversed-phase chromatography as described, it is possible to process approximately 500 mg crude peptide.

The amino acid composition of the synthetic peptide typically is determined following acid hydrolysis at 110° for 24 hr in 200 μl 6 *M* HCl containing 0.1% phenol (v/v) and 1% thioglycol (v/v) to maximize recovery of tyrosine and methionine, respectively. The samples are dried under vacuum, dissolved in 0.2 *M* sodium citrate buffer, and run on a Beckman 6300 automated amino acid analyzer. For a purified synthetic peptide, molar ratios are typically within 2 to 5% of unity. If the method of hydrolysis and amino acid analysis is less precise than this, it has doubtful value as a quality control criteria for peptide purity. Amino acid analysis does not reveal the presence of residual protective groups since these are removed by the acid hydrolysis procedure. However, the presence of most

protective groups is readily detected from the UV spectrum and generally result in distinct reversed-phase chromatography. Capillary electrophoresis[9] is likely to become a more widely used criterion of purity for synthetic peptides since this effectively complements analytical reversed-phase HPLC.

Measurement of Peptide Phosphorylation

There have been a variety of protein kinase assays developed that utilize synthetic peptide substrates. These include spectrophotometric,[10] fluorescent,[11] and radioisotopic methods.[12] Quantitation of the transfer of $^{32}P_i$ to the peptide requires a simple separation system to remove [γ-^{32}P]ATP. The most widely used method has been the P81 phosphocellulose cation-exchange paper assay.[12,13] This procedure depends on the presence of multiple basic residues in the synthetic peptide for binding to the P81 paper. For other peptides it is possible to use the anion-exchange column procedure,[14] gel electrophoresis,[15] thin-layer and paper electrophoresis,[16,17] and isoelectric focusing.[18] A tandem column procedure combining both cation- and anion-exchange chromatography has been developed by Egan *et al.*[19] that gives very low backgrounds, but it is restricted to use with basic synthetic peptides. A number of workers have added arginine residues to nonbasic synthetic peptide substrates to adapt them to the P81 paper assay (see [9] in this volume). Both the column separation procedures and the P81 paper assays can employ Cerenkov counting so there is no need for the continued use of scintillation fluids in many cases.

It is important when characterizing a new peptide substrate to measure the stoichiometry of phosphorylation. In addition to providing information about the specificity of the enzyme this also acts as a quality control for the synthetic peptide. The site of phosphorylation can be determined either

[9] W. G. Kuhr, *Anal. Chem.* **62**, 403 (1990).
[10] H. N. Bramson, E. T. Kaiser, and A. S. Mildvan, *CRC Crit. Rev. Biochem.* **15**, 93 (1984).
[11] D. E. Wright, E. S. Noiman, P. B. Chock, and V. Chau, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6028 (1981).
[12] R. Roskoski, Jr, this series, Vol. 99, p. 3.
[13] D. B. Glass, R. A. Masaracchia, J. R. Feramisco, and B. E. Kemp, *Anal. Biochem.* **87**, 566 (1978).
[14] B. E. Kemp, E. Benjamini, and E. G. Krebs, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1038 (1976).
[15] T. Hunter, N. Ling, and J. A. Cooper, *Nature (London)* **311**, 480 (1984).
[16] T. W. Wong and A. R. Goldberg, *J. Biol. Chem.* **258**, 1022 (1983).
[17] T. Hunter, *J. Biol. Chem.* **257**, 4843 (1982).
[18] C. J. Fiol, A. Wang, R. W. Roeske, and P. J. Roach, *J. Biol. Chem.* **265**, 6061 (1990).
[19] J. J. Egan, M. K. Chang, and C. Londos, *Anal. Biochem.* **175**, 552 (1988).

by the direct sequencing procedures,[20] mass spectrometry,[21] or by more classical approaches.[22] The introduction of capillary electrophoresis has also provided an important analytical tool for characterizing phosphopeptides[9] due to its great sensitivity and resolving power with polar peptides (Applied Biosystems, bioseparations application notes).

Design of Peptide Substrates

There are three approaches to the design of protein kinase peptide substrates. The most widely used approach is to synthesize analogs of known phosphorylation site sequences. These may be from natural substrates, autophosphorylation sites, or phosphorylation sites in exogenous substrates such as histone, myelin basic protein, or caseins. The second approach is to use random polymers of amino acids such as Tyr and Glu for tyrosine kinases (see [7] in this volume). While random polymers may be excellent substrates the disadvantage of this approach is that the phosphorylation site is less well defined. In principle, degenerate random peptide sequences could be used in a cocktail to study the specificity of a protein kinase. The most rapidly phosphorylated peptides could then be isolated and their sequence determined. It would be possible to optimize the kinetic properties of a given peptide substrate by "shotgun" substitution of all other amino acids at every residue. Analogous approaches have already been used to study antibody specificity and to epitope map proteins. Geysen *et al.*[23] and Houghton[24] have developed procedures for making small amounts of large numbers of peptides that could be employed for studying protein kinase specificity in this way. The third approach is to prepare substrate analogs of the pseudosubstrate autoregulatory regions (see below) that have been found in a number of protein kinases. A report of engineering phosphorylation sites into recombinant proteins[25] has been made; however, this is beyond the scope of this chapter.

A number of attempts have also been made to understand phosphoryla-

[20] R. E. H. Wettenhall, R. H. Aebersold, L. E. Hood, and S. B. H. Kent, this series, Vol. 201 [15].
[21] H. E. Meyer, E. Hoffmann-Posorske, H. Korte, and L. M. Heilmeyer, *FEBS Lett.* **204**, 61 (1986).
[22] R. B. Pearson, R. Jakes, M. John, J. Kendrick-Jones, and B. E. Kemp, *FEBS Lett.* **168**, 108 (1984).
[23] H. M. Geysen, R. H. Meloen, and S. J. Barteling, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3998 (1984).
[24] R. A. Houghton, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5131 (1985).
[25] K. Nakai and M. Kanehisa, *J. Biochem. (Tokyo)* **104**, 693 (1988).

tion site sequences in terms of their possible secondary structures.[25-29] Although it is not yet possible to design potent synthetic peptides purely from theoretical structural considerations alone, the rapid expansion in the phosphorylation site sequence database, improved structure predictions, and increased knowledge of protein kinase specificity may make this possible in the future. An alternative to using synthetic peptide substrates directly is to conjugate them to carrier proteins or even immobilize them on various resins. Little has been done in this area except that tandem repeat peptide substrates have been used with the cell cycle kinase,[30] cdc2, and peptide substrates have been conjugated to antibodies as a means of labeling the antibody. It is possible that the kinetics of peptide phosphorylation may be enhanced by conjugation to a carrier protein. A difficulty with chemically conjugated peptide substrates is that, depending on the strategy used, the product may be heterogenous. Peptide substrates can be conjugated to proteins or solid supports through cysteine residues at the carboxyl- or amino-terminal end without interfering with side chains required for recognition by the protein kinase. It is desirable to use a heterobifunctional maleimide cross-linking reagent to give a thioether linkage that is stable in the presence of sulfhydryl reagents. A wide variety of chemical cross-linking procedures have been used to conjugate peptides to proteins to enhance their immunogenicity and some of these could be applied to conjugation and immobilization of protein kinase synthetic peptide substrates (see this series).

A list of synthetic peptide substrates for a number of protein kinases prepared from the local phosphorylation site sequences of known substrates is given in Table I.[31-63] The degree of specificity varies widely with

[25a] B. L. Li, J. A. Langer, B. Schwartz, and S. Pestka, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 558 (1989).
[26] R. E. Williams, *Science* **192**, 473 (1976).
[27] J. F. Leszczynski and G. D. Rose, *Science* **234**, 849 (1986).
[28] C. Radziejewski, W. T. Miller, S. Mobashery, A. R. Goldberg, and E. T. Kaiser, *Biochemistry* **28**, 9047 (1989).
[29] D. A. Tinker, E. G. Krebs, I. C. Felthan, S. K. Attah-Poku, and V. S. Ananthanarayanan, *J. Biol. Chem.* **263**, 5024 (1988).
[30] L. J. Cisek and J. L. Corden, *Nature (London)* **339**, 679 (1989).
[31] D. B. Glass and E. G. Krebs, *J. Biol. Chem.* **254**, 9728 (1979).
[32] R. B. Pearson, J. R. Woodgett, P. Cohen, and B. E. Kemp, *J. Biol. Chem.* **260**, 14471 (1985).
[33] Y. Hashimoto and T. R. Soderling, *Arch. Biochem. Biophys.* **252**, 418 (1987).
[34] B. E. Kemp and R. B. Pearson, *J. Biol. Chem.* **260**, 3355 (1985).
[35] C. H. Michnoff, B. E. Kemp, and J. T. Stull, *J. Biol. Chem.* **261**, 8320 (1986).
[36] B. E. Kemp and M. John, *Cold Spring Harbor Conf. Cell Proliferation* **8**, 331 (1980).
[37] C. House, R. E. H. Wettenhall, and B. E. Kemp, *J. Biol. Chem.* **262**, 772 (1987).
[38] I. Yasuda, A. Kishimoto, S. Tanaka, M. Tominaga, A. Sakurai, and Y. Nishizuka, *Biochem. Biophys. Res. Commun.* **166**, 1220 (1990).

a number of the peptides acting as substrates for multiple protein kinases. In the case of the glycogen synthase kinase-3 substrate peptide listed, prior phosphorylation with casein kinase II is required[51] before it can act as a substrate because of the specificity requirement for an $n + 4$ acidic residue which can be provided by phosphorylated serine [Ser(P)] in this case. In some cases multiple synthetic peptide substrates have been prepared by different laboratories for particular protein kinases. Generally only the one with the most favorable kinetics of phosphorylation has been listed. The protein kinase C peptide substrate PLSRTLS*VAAKK derived from glycogen synthase has the most favorable kinetics but the peptide QKRPS*QRSKYL derived from myelin basic protein is probably more

[39] J. R. Woodgett, K. L. Gould, and T. Hunter, *Eur. J. Biochem.* **161,** 177 (1986).
[40] A. D. Blake, R. A. Mumford, H. V. Strout, E. E. Slater, and C. D. Strader, *Biochem. Biophys. Res. Commun.* **147,** 168 (1987).
[41] B. Gabrielli, R. E. H. Wettenhall, B. E. Kemp, M. Quinn, and L. Bizonova, *FEBS Lett.* **175,** 219 (1984).
[42] S. L. Pelech, B. B. Olwin, and E. G. Krebs, *Proc. Natl. Acad. Sci. U.S.A.* **83,** 5968 (1986).
[43] S. P. Davies, D. Carling, and D. G. Hardie, *Eur. J. Biochem.* **186,** 123 (1989).
[44] P. Hohmann and R. S. Greene, *Biochem. Biophys. Res. Commun.* **168,** 763 (1990).
[45] M. Peter, J. Nakagawa, M. Dorée, J. C. Labbé, and E. A. Nigg, *Cell (Cambridge, Mass.)* **60,** 791 (1990).
[46] P. R. Vulliet, F. L. Hall, J. P. Mitchell, and D. G. Hardie, *J. Biol. Chem.* **264,** 16292 (1989).
[47] J. L. Countaway, I. C. Northwood, and R. J. Davis, *J. Biol. Chem.* **264,** 10828 (1989).
[48] F. Meggio, J. W. Perich, H. E. Meyer, E. Hoffmann-Posorske, D. P. W. Lennon, R. B. Johns, and L. A. Pinna, *Eur. J. Biochem.* **186,** 459 (1989).
[49] P. Agostinis, L. A. Pinna, F. Meggio, O. Marin, J. Goris, J. R. Vandenheede, and W. Merlevede, *FEBS Lett.* **259,** 75 (1989).
[50] E. A. Kuenzel, J. A. Mulligan, J. Sommercorn, and E. G. Krebs, *J. Biol. Chem.* **262,** 9136 (1987).
[51] C. J. Fiol, A. M. Mehrenholz, Y. Wang, R. W. Roeske, and P. J. Roach, *J. Biol. Chem.* **262,** 14042 (1987).
[52] K. J. Chan, *Biochem. Biophys. Res. Commun.* **165,** 93 (1989).
[53] R. A. Masaracchia, B. E. Kemp, and D. A. Walsh, *J. Biol. Chem.* **252,** 7109 (1977).
[54] K. Palczewski, A. Arendt, J. H. McDowell, and P. A. Hargrave, *Biochemistry* **28,** 8764 (1989).
[55] M. H. Watson, A. K. Tancja, R. S. Hodges, and A. S. Mak, *Biochemistry* **27,** 4506 (1988).
[56] G. S. Baldwin, A. W. Burgess, and B. E. Kemp, *Biochem. Biophys. Res. Commun.* **109,** 656 (1982).
[57] C. House, G. S. Baldwin, and B. E. Kemp, *Eur. J. Biochem.* **140,** 363 (1984).
[58] W. Weber, P. J. Bertics, and G. Gill, *J. Biol. Chem.* **259,** 14631 (1984).
[59] J. E. Casnellie and E. G. Krebs, *Adv. Enzyme Regul.* **22,** 501 (1984).
[60] M. L. Harrison, P. S. Low, and R. L. Geahlen, *J. Biol. Chem.* **259,** 9248 (1984).
[61] L. A. Stadtmauer and O. M. Rosen, *J. Biol. Chem.* **258,** 6682 (1983).
[62] J. G. Foulkes, B. Mathey-Prévot, B. C. Guild, R. Prynes, and D. Baltimore, *in* "Cancer Cells/3, Growth Factors and Transformation" (J. Feramisco, B. Ozanne, and C. Stiles, eds.), p. 319. Cold Spring Harbor Lab., Cold Spring Harbor, New York, 1985.
[63] J. A. Cooper, F. S. Esch, S. S. Taylor, and T. Hunter, *J. Biol. Chem.* **259,** 7835 (1984).

TABLE I

PEPTIDE SUBSTRATES BASED ON PROTEIN PHOSPHORYLATION SITE ANALOGS IN NATURAL AND EXOGENOUS SUBSTRATES

Enzyme	Peptide sequence[a]	K_m (μM)	V_{max} (μmol/min/mg)	Ref.[b]
Serine/threonine kinases				
cAMP-PK (mammalian)	LRRAS*LG	4.5	16	1
cAMP-PK (yeast)	VKRKYLKKLTRRAS*FSAQ	4.3	28	*c*
cGMP-PK	MDKVQYLTRSAIRRAS*TIE-MPQQARQNLQNL	7.0	5	*d*
cGMP-PK	RKRS*RAE	29	20	31
Cam II PK[e]	PLRRTLS*VAA	3.5	11.3	32
	PLARTLS*VAGLPGKK	12	2.75	33
Sm MLCK	KKRAARATS*NVFA	7.5	1.4	34
Sk. MLCK	AKRAARATS*NVFS	10	31	35
Phosphorylase kinase	KAKQIS*VRGSL	900	2.9	36
Protein kinase C (mixed)	PLSRTLS*VAAKK	4	12	37
	QKRPS*QRSKYL	7	d	38
	AKRRRLSS*LRA	0.51	1.1	37
	VRKRT*LRRL	48	0.99	39
	YQRRQRKS*RRTI	24	0.92	39
	myr-GSSKSKPKDPS*QRRRSLE	48	0.56	39
	CNle-RRSSSKAYG	4.1	5.0	40
S6 kinase	RRLSS*LRA	180	*f*	41, 42
AMP-PK	HMRSAMS*GLHLVKRR	30	1	43
$p34^{cdc2}$	KS*PAKT*PVK	*f*	*f*	44
	$(S*PTS*PSY)_6$	200	*f*	30
	AVT*PAKKAAT*PAKKA	20[g]	*f*	45
Proline-dependent PK	PTPSAPS*PQPKG	~50	*f*	46
Growth factor PK (p45)	ELVEPLT*PSGEAPNQALLR	*f*	*f*	47
Casein kinase I	DDDEES*ITRR	1000	*h*	48
Casein kinase II	ESLS*SSEE-NHMe	11	*f*	49
	RRREEES*EEE	180	2.13	50
	RRRDDDS*DDD	60	2.19	50
Mammary gland PK	ESLSS*SEE-NHMe	38	*f*	49
Glycogen synthase kinase-3	PRPAS*VPPS*PSLS*RH SS*PHQSEDEEEP	2	*f*	51
Ganglioside PK	RFS*WGAEGQK	*f*	*f*	52
Histone H_4 kinase I	VKRIS*GLG	43	0.016	53
Rhodopsin kinase	-DEASTTVSKTETSQVAP-	1400	0.008	54
Tropomyosin kinase	KLKYKAISEELDHALNDMTS*I	500	0.034	55
Protein-tyrosine kinases				
EGF-receptor kinase	RRLEEEEEAY*G	150	0.002	56
	LIEDAEYTA	440	0.006	57
	DRVY*IHPF (angiotensin II)	800	0.011	58
	RREELQDDY*EDD	90	0.001	*i*

(*continued*)

TABLE I (*continued*)

Enzyme	Peptide sequence[a]	K_m (μM)	V_{max} (μmol/min/mg)	Ref.[b]
$p56^{lck}$	RRLIADAEY*AARG	1300	0.006	59
	DRVY*IHPFHL (angiotensin I)	2300	0.005	60
$pp60^{v\text{-}src}$	DRVY*IHPFHL (angiotensin I)	1540	0.003	16
	DRVY*IHPF (angiotensin II)	2000	0.010	16
	DRVY*VHPF [V^5]angiotensin II)	870	0.007	16
	IENEEQEY*VQTVK	440	0.010	28
	Raytide	100	*f*	*j*
Insulin receptor	RRLIEDAEY*ARG	1200	*f*	61
	DRVY*IHPF (angiotensin II)	2600	*f*	61
	RVY*VHPF (angiotensin III inhibitor)	8000	*f*	61
$pt^{abl}50$	DRVY*IHPFHL (angiotensin I)	3700	1.25	62
$p140^{gag\text{-}fps}$	AAVPSGASTGIY*EALELR (enolase peptide)	200	*f*	63

[a] *, Phosphate acceptor site.
[b] Numbers refer to text footnotes.
[c] C. Denis *et al.* (unpublished, 1990); J. R. Cherry, T. R. Johnson, C. Dollard, J. R. Shuster, and C. L. Denis, *Cell* (*Cambridge, Mass.*) **56**, 409 (1989).
[d] P. J. Robinson, B. Michell, K. I. Mitchelhill, and B. E. Kemp (unpublished, 1990).
[e] Cam, calmodulin; Sm, smooth; Sk, skeletal; MLCK, myosin light chain kinase; myr, myristate; Nle, norleucine; Me, methylester.
[f] Not reported or not given per unit protein.
[g] Inferred from K_i value.
[h] V_{max} equivalent to β-casein A^2.
[i] C. House and B. E. Kemp (erythrocyte band 3 site, unpublished).
[j] Raytide is a model peptide substrate supplied by Oncogene Sciences, Inc. Although the structure of this peptide is not provided by the manufacturer, it has a K_m value similar to the gastrin and band 3 peptides (see above).

specific (Table I). The data concerned with protein kinase C refers to the preparation from brain consisting of a mixture of α, β, and γ isoenzymes. There are not yet any isoenzyme-specific synthetic peptide substrates available for protein kinase C, but work in this area is being undertaken.[64] The S6 kinase refers to the growth factor-sensitive enzyme activity; how-

[64] R. M. Marais and P. J. Parker, *Eur. J. Biochem.* **182**, 129 (1989).

TABLE II
PEPTIDE SUBSTRATES BASED N AUTOPHOSPHORYLATION SITES

Enzyme	Peptide	K_m (μM)	V_{max} (μmol/min/mg)	Ref.[a]
Cam II-PK	MHRQET*VDC	10	3.15	66
cGMP-PK	IGPRTT*RAQGI	578	0.069	67
EGF-receptor PK	RRKGSTAENAEY*LRV	160	0.009	68
	RRISKDNPDY*QQD	340	0.010	68
	RRDDTFLPVPEY*INQS	410	0.011	68
Insulin-receptor PK	TRDIY*ETDY*Y*RK	240	ND[b]	69
$pp60^{src}$	EDNEY*TARQG	6250	0.001	16
$pp60^{src}$	EDNEY*VARQG	5890	0.001	16
$p56^{lck}$	PRLIEDAEY*AARG	1160	0.01	29

[a] Numbers refer to text footnotes and these contain the names of the parent substrate proteins.
[b] ND, Not reported.

ever, it is now recognized that there are multiple S6 kinases. There are a number of protein kinases capable of phosphorylating Thr/Ser-Pro sites in addition to $p34^{cdc2}$ and further work is required to understand their comparative specificities. A detailed account of tyrosine kinase synthetic peptide substrates has recently been prepared by Geahlen and Harrison.[65] For several protein kinases consensus phosphorylation site sequences are known, such as RRXS*X for the cyclic AMP-dependent protein kinase and RXXS*XR for protein kinase C. The corresponding synthetic peptides tend to be good substrates, but it should be noted that the consensus sequence is usually recognized by comparing multiple phosphorylation site sequences as well as taking into account peptide substrate structure/function data in which the roles of key specificity determinants have been studied.

Many protein kinases undergo autophosphorylation and there are a number of examples where autophosphorylation site sequences have been used to construct synthetic peptide substrates (Table II).[66–69] The kinetics of phosphorylation of these peptides is generally no better, and frequently worse, than for peptides based on the local phosphorylation site sequences

[65] R. J. Geahlen and M. L. Harrison, *in* "Peptides and Protein Phosphorylation" (B. E. Kemp, ed.), p. 239. Uniscience CRC Press, Boca Raton, Florida, 1990.
[66] R. J. Colbran, Y. L. Fong, C. M. Schworer, and T. R. Soderling, *J. Biol. Chem.* **263**, 18145 (1988).
[67] D. B. Glass and S. B. Smith, *J. Biol. Chem.* **258**, 14797 (1983).
[68] J. Downward, M. D. Waterfield, and P. J. Parker, *J. Biol. Chem.* **260**, 14538 (1985).
[69] L. Stadtmauer and O. M. Rosen, *J. Biol. Chem.* **261**, 10000 (1986).

TABLE III
PEPTIDE SUBSTRATES BASED ON PSEUDOSUBSTRATE PROTOTOPES

Enzyme	Peptide	K_m (μM)	V_{max} (μmol/min/mg)	Ref.[a]
αPK-C(19–31), S_{25}	RFARKGS*LRQKNV	0.2	8	70
αPK-C(15–31), S_{25}	DVANRFARKGS*LRQKNV	18	4.5	64
β_1PK-C(15–31), S_{25}	ESTVRFARKGS*LRQKNV	7.2	1.9	64
γPK-C(15–31), S_{25}	GPRPLFCRKGS*LRQKVV	9.6	2.3	64
εPK-C(149–164), S_{159}	ERMRPRKRQGS*VRRRV	68	1.9	71
PKI(14–22), S_{22}	GRTGRRNS*I	0.11	9.2	72

[a] Numbers refer to text footnotes.

of substrates. Accordingly none of these is used routinely as model substrates.

Potent synthetic peptide substrates have been constructed using the pseudosubstrate sequences found in protein kinase C and the cAMP-dependent protein kinase inhibitor protein. Some of these peptides have K_m values in the submicromolar range (Table III).[70–72] On the other hand, the pseudosubstrate-based peptide substrates for the myosin light chain kinases from both smooth muscle and skeletal muscle are exceedingly poor substrates (not listed). Substrate analogs of the pseudosubstrate regions have been prepared for protein kinase C isoenzyme forms α, β_1, and γ and all act as effective substrates (Table III), but are not isoenzyme specific.

Schaap and Parker[71] exploited the pseudosubstrate idea to design a peptide substrate for recombinant ε protein kinase C that does not phosphorylate histone. As more protein kinases are identified by cDNA cloning techniques, particularly through the use of the polymerase chain reaction, the demand for peptide substrates designed using the pseudosubstrate regulatory concept will increase.

The specificity of synthetic peptide substrates toward different protein kinases may be modulated by making amino acid substitutions. An early example of this was shown for the phosphorylase peptide,[2,36] KRKQIS$_{14}$VRGL, where replacement of Arg-16 with Ala made the peptide a potent substrate for the cAMP-dependent protein kinase and a poor

[70] C. House and B. E. Kemp, *Science* **238**, 1726 (1987).
[71] D. Schaap and P. J. Parker, *J. Biol. Chem.* **265**, 7301 (1990).
[72] D. B. Glass, H. C. Cheng, L. Mende-Mueller, J. Reed, and D. A. Walsh, *J. Biol. Chem.* **264**, 8802 (1989).

TABLE IV
MODULATION OF PEPTIDE SUBSTRATE SELECTIVITY BY SUBSTITUTION

Peptide sequence [Phosphorylase(9–18)[a]]	Phosphorylase kinase K_m (μM)	Phosphorylase kinase V_{max} (μmol/min/mg)	cAMP-PK K_m (μM)	cAMP-PK V_{max} (μmol/min/mg)	Ref.[a]
KRKQIS*VRGL	900	2.9	3900	4.1	70
KRKQIS*VAGL	2500	0.18	36	21.4	64
KAKQIS*VRGL	900	2.7	2200	0.04	64
					64
					71
					72

[a] From Ref. 2.

substrate for phosphorylase kinase (Table IV). Substitution of Arg-10 with Ala had the opposite effect, enhancing specificity toward phosphorylase kinase and suppressing phosphorylation by the cAMP-dependent protein kinase. Similar switches in specificity have been engineered for the glycogen synthase peptide,[32] PLSRTLS$_7$VAA, where substitution with Arg instead of Ser-3 favors phosphorylation with the calmodulin-dependent kinase II. Substitution of Thr-5 with Arg, however, favors phosphorylation by the cAMP-dependent protein kinase.

Applications

Synthetic peptide substrates have proved extremely useful reagents in the study of protein kinases across a wide spectrum of studies, extending from those carried out with crude extracts to structural studies with NMR and X-ray crystallography. By using peptides containing a single phosphorylatable residue it has been possible to detect protein kinases in cell extracts containing multiple protein kinase activities. The sensitivity and specificity of synthetic peptide substrates has made them the substrate of choice in the study of hormonal regulation of protein kinases.[73] Stability and chemical purity are also major benefits in using synthetic peptide substrates. Since quantities are not usually limiting they can often be used at saturating concentrations, making it possible to obtain maximum rates of phosphorylation. Synthetic peptides may be used effectively in the confirmation of natural phosphorylation sites. Often the available protein substrate is limiting; synthesis of peptides corresponding to a number of potential phosphorylation sites in the protein can be used to determine the

[73] S. A. Livesey, B. E. Kemp, C. A. Re, N. C. Partridge, and T. J. Martin, *J. Biol. Chem.* **257**, 14983 (1982).

most likely site.[74] The power of this approach can be further enhanced by making complementary point mutations in the protein substrate. With improvements in the potency of peptide substrates it is likely that applications of affinity purification will also increase ([13], this volume).

In the past the development of synthetic peptide substrates has lagged behind the discovery of protein kinases. This situation is changing due to the widespread use of many synthetic peptides, such as the kemptide and the ribosomal S6 peptide, which have been used to detect protein kinases with unforeseen overlapping specificities to the cAMP-dependent protein kinase.[75,76] The availability of relatively specific inhibitors such as the cyclic AMP-dependent protein kinase inhibitor peptide PKI(5–22)[77] and other pseudosubstrate inhibitors (this series, Volume, 201 [24]) as well as calcium chelators for calcium-dependent protein kinases greatly facilitates attempts to detect new protein kinase activities using peptide substrates capable of being phosphorylated by multiple protein kinases.

It can be expected that in the forthcoming years there will be even greater synergy between the use of recombinant protein expression and synthetic peptides to create a variety of protein kinase substrates in order to explore the mechanisms of regulation by protein phosphorylation.

[74] B. Luscher, E. Christenson, D. W. Litchfield, E. G. Krebs, and R. N. Eisenman, *Nature (London)* **344**, 517 (1990).
[75] E. Erikson and J. L. Maller, *Second Messengers Phosphoproteins* **12**, 135 (1988).
[76] J. K. Klarlund, A. P. Bradford, M. G. Milla, and M. P. Czech, *J. Biol. Chem.* **265**, 227 (1990).
[77] B. E. Kemp, H. C. Cheng, and D. A. Walsh, this series, Vol. 159, p. 173.

[11] Synthetic Peptide Substrates for Casein Kinase II

By DANIEL R. MARSHAK and DENNIS CARROLL

Casein Kinase II

The enzyme casein kinase II (CKII) is a protein-serine/threonine kinase found in all eukaryotic cells.[1] Its ubiquitous distribution among species and tissues implies a function central to all nucleated cells. CKII was first identified from rabbit reticulocyte lysates,[2] and subsequently isolated from hypotonic extracts of mammalian liver and lung tissue.[3] Following

[1] G. M. Hathaway and J. A. Traugh, *Curr. Top. Cell. Regul.* **21**, 101 (1982).
[2] G. M. Hathaway and J. A. Traugh, *J. Biol. Chem.* **254**, 762 (1979).

EXHIBIT 3

Beebe for comments on the manuscript. We acknowledge partial support for studies reviewed here from the US-Spain Joint Committee for Scientific and Technological Cooperation.

References

1 Ham, R. G. and Veomett, M. J. (1980) in *Mechanisms of Development*, p. 429 The C. V. Mosby Company
2 Tata, J. R. (1984) in *Biological Regulation and Development*, Vol. 3B (Goldberger, R. P. and Yamamoto, K. R., eds), p. 13, Plenum Press
3 Gilbert, S. F. (1988) *Developmental Biology*, 2nd Edn, Sinauer
4 Gudernatsch, J. F. (1912) *Wilhelm Roux's Arch. Entwicklungsmech. Org.* 35, 457–483
5 Kopec, S. (1922) *Biol. Bull.* (Woods Hole, MA) 24, 323–342
6 Levi-Montalcini, R. and Angeletti, P. U. (1969) *Physiol Rev.* 48, 534–568
7 Cohen, S. and Savage, C. R., Jr (1974) *Recent Prog. Horm. Res.* 30, 551–574
8 De Pablo, F., Roth, J., Hernandez, E. and Pruss, R. M. (1982) *Endocrinology* 111, 1909–1916
9 Serrano, J., Bevins, C. L., Young, W. S. and De Pablo, F. (1989) *Dev. Biol.* 132, 410–418
10 Girbau M., Bassas, L., Alemany, J. and De Pablo, F. (1989) *Proc. Natl Acad. Sci. USA* 86, 5868–5872
11 De Pablo, F., Girbau, M., Gomez, J. A., Hernandez, E. and Roth, J. (1985) *Diabetes* 34, 1063–1067
12 Girbau, M., Lesniak, M. A., Gomez, J. A. and De Pablo, F. (1988) *Biochem. Biophys. Res. Commun.* 153, 142–148
13 Scavo, L., Alemany, J., Roth, J., and De Pablo, F. (1989) *Biochem. Biophys. Res. Commun.* 162, 1167–1173
14 Beebe, D. C., Silver, M. H., Belcher, K. S., Van Wyk, J. J., Svoboda, M. E. and Zelenka, P. S. (1987) *Proc. Natl Acad. Sci. USA* 84, 2327–2330
15 Alemany, J., Zelenka, P., Serrano J. and De Pablo F. (1989) *J. Biol. Chem.* 264, 17559–17563
16 Spemann H. and Mangold H. (1924) *Wilhelm Roux's Arch. Entwicklungsmech. Org.* 100, 599–638
17 Slack, J. M. W. (1983) in *From Egg to Embryo*, p. 26, Cambridge University Press
18 Nieuwkoop, P. D. (1969) *Wilhelm Roux's Arch. Entwicklungsmech. Org.* 162, 341–373
19 Sargent, T. D., Jamrich, M. and David, I. B. (1986) *Dev. Biol.* 114, 238–246
20 Gurdon, J. B., Brennan, S., Fairman, S. and Mohun, T. J. (1984) *Cell* 38, 691–700
21 Weeks, D. L. and Melton, D. A. (1987) *Cell* 51, 861–867
22 Kimelman, D. and Kirschner, M. (1987) *Cell* 51, 869–877
23 Rosa, F., Roberts, A. B., Danielpour, D., Dart, L. L., Sporn, M. B. and Dawid, I. B. (1988) *Science* 239, 783–785
24 Kimelman, D., Abraham, J. A., Haaparanta, T., Palisi, T. M. and Kirschner, M. W. (1988) *Science* 242, 1053–1056
25 Gillespie, L. L., Paterno, G. D. and Slack, J. M. W. (1989) *Development* 106, 203–208
26 Smith, J. C., Price, B. M. J., Van Nimmen, K. and Huylebroeck, D. (1990) *Nature* 345, 729–731
27 Rosen, O. M. (1987) *Science* 237, 1452–1458
28 Livneh, E., Glazer, L., Segal, D., Schlessinger, J. and Shilo, B-Z. (1985) *Cell* 40, 599–607
29 Miodzik, M., Hiromi, Y., Weber, U., Goodman, C. S. and Rubin, G. M. (1990) *Cell* 60, 211–224
30 Vassin, H., Bremer, K. A., Knust, E. and Campos-Ortega, J. (1987) *EMBO J.* 6, 3431–3440
31 Greenwald, I. (1985) *Cell* 43, 583–590
32 Yang, Q., Angerer, L. M. and Angerer, R. C. (1989) *Science* 246, 806–808
33 Wright, C. V. E., Cho, K. W. Y., Oliver, G. and De Robertis, E. M. (1989) *Trends Biochem Sci.* 14, 52–56
34 Ruiz I Altaba, A. and Melton, D. A. (1990) *Trends Genet.* 6, 57–64
35 Rappolee, D. A., Brenner, C. A., Schultz, R., Mark, D and Werb, Z. (1988) *Science* 241, 1823–1825
36 Mattson, B. A., Rosenblum, I. Y., Smith, R. M. and Heyner, S. (1988) *Diabetes* 37, 585–589
37 Serrano, J., Shuldiner, A. R., Roberts, C. T., LeRoith, D. and De Pablo, F. *Endocrinology* (in press)
38 Herrmann, B. G., Labeit, S., Poustka, A., King, T. R and Lehrach, H. (1990) *Nature* 343, 617–622
39 Hamburger, V. and Hamilton, H. L. (1951) *J. Morphol.* 88, 49–92

Protein sequence MOTIFS

Protein kinase recognition sequence motifs

Bruce E. Kemp and Richard B. Pearson

Protein kinases play a crucial role in the regulation of many cellular processes. They alter the functions of their target proteins by phosphorylating specific serine, threonine and tyrosine residues. Identification of phosphorylation site sequences and studies with corresponding model peptides have provided clues to how these important enzymes recognize their substrate proteins. This knowledge has made it possible to identify potential sites of phosphorylation in newly sequenced proteins as well as to construct specific model substrates and inhibitors.

A REVIEW of protein kinase recognition sequences is both long overdue and premature. It is overdue because many of the approaches and applications of this field are more than a decade old, and it is premature because we do not know the three-dimensional structure of a single protein kinase substrate complex. The study of protein kinases over the last 35 years has resulted in protein phosphorylation being recognized as one of the most important mechanisms of regulating intracellular processes. There are few, if any, physiological processes in eukaryotes that are not dependent on protein phosphorylation. While this brief review is focused on protein kinase recognition motifs, it should be recognized that protein phosphatases, which catalyse the reverse reaction, are equally important players in the overall process of regulation of protein function by phosphorylation. Undoubtedly, their specificity and regulatory properties are no less important.

B. E. Kemp and **R. B. Pearson** are at the St Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia.

All protein kinases contain a common catalytic domain which typically extends over 240 residues[1], including the binding sites for ATP and the protein substrate (Fig. 1). The ATP-binding site is located at the amino terminus of the domain as characterized by the Rossmann motif, GXGXXG, while the centrally located Asp184 is responsible for base catalysed transfer of the phosphate to the protein substrate[2]. The binding site for the protein substrate is not unequivocally established, although some evidence favours the idea that it is located in the carboxy-terminal 60 residues of the catalytic domain[3].

Early studies found that protein kinases phosphorylated their target proteins at discrete sites. These enzymes were shown to prefer certain exogenous substrates (such as casein, phosvitin and histones) and only phosphorylated a limited number of available sites. Indeed, Krebs and Fischer showed that phosphorylase kinase only phosphorylated Ser14 in phosphoryl-

 0376–5067/90/$02.00

ase which contains 64 Ser and Thr residues. Langan's pioneering work on histone phosphorylation demonstrated that several protein kinases may phosphorylate a common substrate at multiple sites. The importance of substrate specificity was further highlighted by the discovery of the cAMP-dependent protein kinase by Walsh and his colleagues. This enzyme did not phosphorylate phosphorylase but did phosphorylate a wide range of exogenous proteins. The idea that the multiple functions of the second messenger cAMP were all mediated by the cAMP-dependent protein kinase implied the enzyme recognized a panel of protein substrates that were subject to hormonal regulation. This raised the question of how protein kinases recognized specific residues out of the numerous hydroxyl groups in their substrates.

It became evident that the local sequence around the phosphorylation site played a vital role in recognition by the cAMP-dependent protein kinase and that arginine residues were involved. The cAMP-dependent protein kinase phosphorylated β casein B at Ser124 in the sequence TERQSLT (nos 120–126) but not in the more common variant β casein A[2] where Ser replaces Arg at position 122 (Ref. 4). This observation suggested that the cAMP-dependent protein kinase substrate recognition motif may be the RXS* sequence†. Studies in several laboratories in the mid-1970s demonstrated that the cAMP-dependent protein kinase readily phosphorylated short synthetic peptides[5] and provided compelling evidence for the role of Arg residues. Significantly, the synthetic peptide modelled on the liver pyruvate kinase phosphorylation site sequence, LRRASLG (Kemptide), was phosphorylated with kinetic constants comparable to native protein substrates. This data suggests that all of the information necessary for recognition by the cAMP-dependent protein kinase could be present in the local phosphorylation site sequence and that RRXS was a preferred motif (reviewed in Ref. 5). While the primary sequence and proximity of Arg residues plays an important role, higher orders of structure can have an overriding influence.

†The phosphate acceptor site is indicated S* to distinguish it from serine phosphate (S(P)) which may act as a specificity determinant for some protein kinases. Where the specificity determinants are known, less essential residues are marked X and determinant residues are shown in bold. The complete phosphorylation site sequence is given where the major determinants are not known (see Table I).

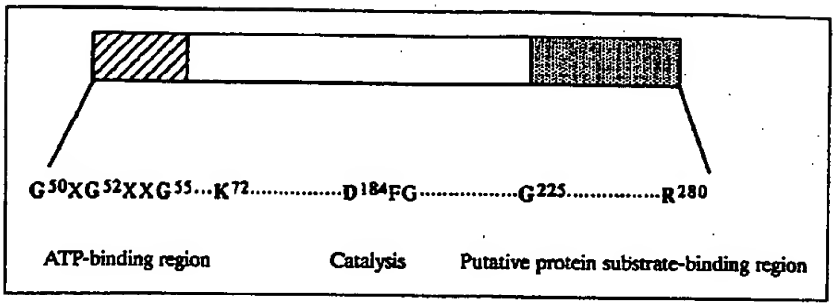


Figure 1
General structure of a protein kinase catalytic domain. Residues conserved in almost all known serine, threonine and tyrosine protein kinases are numbered according to the cAMP-dependent protein kinase catalytic subunit.

For example, lysozyme does not act as a substrate for the cAMP-dependent protein kinase unless it is chemically modified.

Protein kinase specificities

The study of sequences surrounding the local phosphorylation site and the phosphorylation of model peptides has enabled the identification of phosphorylation site motifs for a number of protein kinases (see Table I). Although this is an extensive list, the explosion of protein kinase sequences generated by DNA cloning has outstripped our knowledge of their substrate specificities and identification of natural substrates.

All of the calmodulin-dependent protein kinases studied to date utilize basic residues as specificity determinants. The specificity of phosphorylase kinase was studied in detail by Graves and his colleagues[6]. In the model palindromic peptide LSYRRYSL (nos 1–8), Ser2 is phosphorylated by phosphorylase kinase, whereas the cAMP-dependent protein kinase phosphorylates Ser7 on the carboxyl side of the adjacent arginines[6]. While there is a strong influence of a carboxyl basic residue S*XR in the phosphorylase peptide, this is not an absolute requirement since the enzyme can phosphorylate sites without a basic residue in this position. The non-charged adjacent residues may also influence the phosphorylation of model peptides. The myosin light-chain kinases from both smooth and skeletal muscle also show restricted substrate specificity for myosin light chains. While the skeletal muscle isoenzyme will phosphorylate light chains from skeletal, cardiac and smooth muscle, Stull and his colleagues have shown that the smooth muscle enzyme has a strong preference for light chains from the same muscle. The smooth muscle myosin light-chain kinase requires the sequence KKRXXRXXS*, with the number and spatial arrangement of the basic residues essential for favourable kinetics of phosphorylation and for directing the phosphate to the correct site[7]. In skeletal muscle myosin light chains the local phosphorylation site sequence contains Glu at residue 10 in the sequence PKKAKRRAAEGSS*NVFS (nos 1–17). Synthetic peptide analogs of the native sequence are phosphorylated with low V_{max} values, whereas those containing Arg at position 10, analogous to the smooth muscle light chains, are readily phosphorylated by skeletal muscle enzyme. This is a good example of a negative determinant that is apparent from peptide studies but presumably not accessible in the intact protein. To some extent studying the recognition requirements of protein kinases with highly restricted specificity ranges, such as phosphorylase kinase and the myosin light-chain kinases, is made difficult because of the lack of multiple natural phosphorylation site sequence for comparisons. This is also a problem for the tyrosine kinases. On the other hand, the multifunctional calmodulin-dependent protein kinases have broad specificities and recognize the motif RXXS*X in both proteins and peptide substrates[8]. The specificity requirements of the other members of the calmodulin-dependent protein kinase family are being explored in several laboratories.

Protein kinase C has been the subject of numerous substrate specificity studies. The initial studies were carried out on brain enzyme which consists of multiple isoenzymes. All protein kinase C preparations have a requirement for basic residues but there can be considerable variation in the juxtaposition and choice of Arg over Lys around the

phosphorylation site. Synthetic peptides containing the motif XRXXS*XRX tend to be the best substrates and indeed this arrangement is present in the pseudosubstrate autoregulatory region of the enzyme (see below). With the availability of individual recombinant isoenzymes of protein kinase C, Parker and his colleagues are exploring the basic residue requirements of the individual isoenzymes.

The specificity and recognition sites for the cyclic nucleotide-dependent protein kinases have been extensively studied (reviewed in Ref. 5). For the cAMP-dependent protein kinase, the most typical motif is RRXS*X but RXS*X and KRXXS*X are also encountered. The phosphorylation site sequence RRS* occurs in both cardiac troponin and hormone-sensitive lipase, but limited synthetic peptide studies indicate that the Arg adjacent to the Ser(*) is less important than the second Arg. There is also evidence from synthetic peptide studies that more distal Arg residues on the amino-terminal side may have a positive influence. Indeed the heat-stable inhibitor (Walsh inhibitor) of this enzyme has the pseudosubstrate motif GRTGRRNA*I with Ala occupying the equivalent of the Ser phosphate acceptor site. In scanning sequences for cAMP-dependent protein kinase sites the search pattern is RXS*, then RRXS* or KRXXS*, and not S*XR as the carboxy-terminal basic residue tends to be deleterious. The only known examples of RXS* motifs where X is not Arg are exogenous substrates phosphorylated *in vitro*. The sites phosphorylated *in vivo* by the cAMP-dependent protein kinase all contain multiple adjacent Arg residues in the arrangements RRXS* or RRS*, with two having the latter motif. A hydrophobic residue is often, but not exclusively, found after the Ser. While the yeast cAMP-dependent protein kinase specificity differs in several respects, the fact that it can be complemented by the mammalian enzyme in yeast suggests that the differences do not override the natural function. Recognition site specificity of the cGMP-dependent protein kinase has been studied in detail by Glass and his colleagues[5]. From model peptide studies, there is evidence for a requirement for an Arg located on the carboxy-terminal side of the phosphorylated residue in the sequence S*R or T*R; however this is not an absolute requirement as the enzyme has been shown to phosphorylate sites without this motif. The precise requirements are insufficiently clear to allow cGMP-dependent protein kinase phosphorylation sites to be confidently identified by scanning amino acid sequences alone.

Table I. Protein kinase phosphorylation site motifs

Protein kinase	Recognition motif	Refs
Serine and threonine kinases		
Phosphorylase kinase	KRKQIS*VR	Chan[6]
Myosin light chain kinase (smooth muscle)	XKKRXXRXXS*X	Kemp[5]
Myosin light chain kinase (skeletal muscle)		
Myosin-I heavy chain kinase	KXXS*X or RXT*X	Brzeska[10]
Calmodulin-dependent protein kinase I	NYLRRLS*DSNF	Czernik[11]
Multifunctional calmodulin-dependent protein kinase II	XRXXS*X	Pearson[8]
Calmodulin-dependent protein kinase III	RAGET*RFT*DT*RK	Nairn[12]
cAMP-dependent protein kinase (mammalian) (yeast)	XRRXS*X XRRXS*X	Zetterqvist[5]
cGMP-dependent protein kinase	XS*RX	Glass[5]
Protein kinase C (α, β, γ)	XRXXS*XRX	Graff[13]
S6 kinase II	XRXXS*X	Erikson[14]
dsRNA-dependent kinase pp68	SELS*RR	Colthurst[15]
dsDNA-dependent kinase	PEET*QT*QDQPMEEEE	Lees-Miller[16]
Protease activated kinase I & II	AKRRRLSS*LRA	Wettenhall[9]
Cell cycle kinase cdc-28, MPF	XKS*PX or XKT*PX	Langan[18]
Proline-dependent protein kinase	XS*PX or XT*PX	Vulliet[19]
Growth factor regulated kinase	PLT*PSGEA	Countaway[20]
Casein kinase I	XS(P)XXS*X or XEXXS*X	Pinna[5]
Casein kinase II	XS*XXEX	Pinna[5]
Mammary gland casein kinase	XS*XEX or XS*XS(P)X	Pinna[5]
Glycogen synthase kinase-3	XS*XXXS(P)X	Fiol[21]
AMP-activated protein kinase (acetyl CoA carboxylase kinase) (HMG-CoA reductase kinase) (hormone sensitive lipase kinase)	MRSSMS*GLHL MIHNRS*KINL MRRSVS*EAAL	Hardie[22]

As early as 1970 Ribadeau-Dumas *et al.*[9] made correct predictions about the specificity requirements of the casein kinase from the mammary gland Golgi apparatus (S*XE or S*XS(P)) based on the sequence of the polyphosphorylated region in casein. Casein kinase II specificity has been studied extensively, in particular in the laboratories of Pinna and Krebs. This enzyme has a very widespread distribution and recognizes the motifs S*XXE and S*XXS(P). Not every site conforming to the motif is phosphorylated, casein kinase II phosphorylates Ser17 and not Ser18 in the peptide from β casein A², ESLSSSEE (nos 14–21). On the other hand, the mammary gland enzyme phosphorylates Ser18 and not Ser17 in this sequence. Casein kinase I and II have been shown to participate in hierarchal phosphorylation reactions by several groups. Roach and his colleagues found that prior phosphorylation of glycogen synthase by cAMP-dependent protein kinase at Ser7 caused casein kinase I to phosphorylate Ser10, whereas phosphorylation of glycogen synthase by casein kinase II provides the recognition

Table I. Protein kinase phosphorylation site motifs (continued)

Protein kinase	Recognition motif	Refs
Pyruvate dehydrogenase kinase	S*MSDPGVS*YRYGMGTS*VE	Edelman[23]
Branched chain α-ketoacid dehydrogenase kinase	GHHS*TSDD and SYRS*VDE	Paxton[24]
Heme regulated eIF-2a kinase	LSELS*RR	Kudlicki[25]
Endogenous eIF-4E kinase	KNDKS*KTWQ	Rychlik[26]
Histone H4 kinase I	VKRIS*GLG	Masaracchia[5]
Histone H4 kinase II	AcS*GRGKGG	Masaracchia[5]
Isocitrate dehydrogenase\ kinase (*E. coli*)	GIRS*LNVALR	Thorsness[27]
β-Adrenergic receptor kinase	GYS*S*NGNT*GEQS*G(X)$_{16}$G T*ED(X)$_5$GT*VPS*DNIDS*Q(X)$_3$ S*T*NDS*LL	Hausdorff[28]
Rhodopsin kinase	DEAS*T*T*VKTETS*QVA	Palczewski[29]
Tropomyosin kinase	DNALNDITS*L-COOH	Watson[30]
Tyrosine kinases		
p60$^{v\text{-}src}$	RLIEDNEY*TARQGAK	Geahlen[5]
p56^{lck}	RLIEDNEY*TAREGAK	Geahlen[5]
p40^{thymus}	PEEDGERY*DEDEE	Geahlen[5]
p85$^{gag\text{-}fps}$	REEADGVY*AASGGLR	Geahlen[5]
p90$^{gag\text{-}yes}$	RKIEDNEY*TAREGAK	Geahlen[5]
p120$^{gag\text{-}abl}$	EEKEY*HAE	Geahlen[5]
EGF receptor	TAENAEY*LRVAP	Geahlen[5]
Insulin receptor	TRDIY*ETDY*Y*RK	Geahlen[5]
p75^{liver}	DRVY*VHPF	Geahlen[5]
Spleen tyrosine kinase	EDAEY*AARRRG	Geahlen[5]

[a] R. E. H. Wettenhall and N. Morrice, unpublished.

Ser(P) for glycogen synthase kinase-3. This enzyme recognizes the motif XS*XXXS(P)X. It is of interest that all four of the Ser/Thr kinases studied that utilize Glu or Asp as specificity determinants also recognize Ser(P). There are also examples where phosphorylation at one site by a particular protein kinase can suppress the phosphorylation of a nearby residue by another protein kinase. The nearby phosphorylation sites in the hormone sensitive lipase, MRRSVSEA (nos 560–567) exhibit this behaviour with mutually exclusive phosphorylation[17] by the cAMP-dependent protein kinase (Ser563) and calmodulin-dependent protein kinase II (Ser565). It seems likely that examples of hierarchal phosphorylation will become more frequent as the specificities of additional protein kinase are studied. So far hierarchal phosphorylation has been observed between nearby phosphorylation site sequences, as well as over distances of 25 residues in the case of the regulatory subunit R_s phosphorylation by casein kinase II and glycogen synthase kinase-3; it is possible that even more distant interactions involving higher orders of structure may occur.

The recognition sites for a number of tyrosine protein kinases have been examined. Inspection of the known phosphorylation sites indicates that acidic residues are often located near the tyrosine phosphate acceptor site and their importance has been demonstrated with model peptides in several instances. In general, few natural substrates (excluding autophosphorylation) have been reported for tyrosine protein kinases and this has meant that we have not had the benefit of comparisons. In summarizing our knowledge of tyrosine phosphorylation site motifs, Geahlen and Harrison[5] noted: 'Observations drawn from the study of synthetic peptides have frustrated efforts to clearly define primary structural determinants that are involved in the recognition of substrates by tyrosine kinases. It is perhaps some consolation to investigators of tyrosine kinase substrate specificity that not all peptides containing tyrosine residues are substrates'. Clearly some fresh approaches are required.

Phosphorylation site motifs

The progress in recognizing specific phosphorylation site motifs for many protein kinases has led to the expectation of being able to scan protein sequences and identify phosphorylation sites for given protein kinases. However, considerable caution is required as there are too many exceptions at present to accept the phosphorylation site motifs listed in Table I as 'canons' of recognition. We do not know precisely which nominal specificity determinants actually have corresponding residues in protein kinase active sites and which just favour an optimum conformation. Hopefully this dilemma will be short lived with the imminent solution of the X-ray structure of the cAMP-dependent protein kinase complex with substrate and inhibitor peptides. Phosphorylation site sequence studies such as those made by Cohen and his colleagues have been of enormous value in contributing to our knowledge of protein kinase specificity and the development of the concepts of recognition motifs, but their most enduring value is in answering the difficult question of which sites are actually functionally significant *in vivo*. Structures resembling protein kinase phosphorylation site motifs also play an important role in the regulation of protein kinases. These structures, called pseudosubstrate prototopes, are located in the regulatory domains of protein kinases and are responsible for maintaining protein kinases in inactive

forms[3]. The pseudosubstrate sequences typically contain an alanine in place of the serine or threonine found in the phosphorylation site motifs. For protein kinase C, the pseudosubstrate sequence occurs between residues 19 and 31, RFARKGA*LRQKNV, which resembles the substrate motif RXXS*XR (Table I) with Ala25 in place of the phosphate acceptor site. In this case, activation of protein kinase C by binding of diacylglycerol is thought to induce a conformational change that removes the pseudosubstrate structure from the active site allowing access to substrates. One cannot fail to be impressed with the way nature has utilized the same features responsible for substrate recognition to regulate these important enzymes.

References

Because of the limit on the number of references it has not been possible to cite many significant contributions, but these can be traced from those given.

1 Hanks, S. K., Quinn, A. M. and Hunter, T. (1988) *Science* 241, 42–52
2 Taylor, S. S. (1989) *J. Biol. Chem.* 264, 8443–8446
3 House, C., Robinson, P. J. and Kemp, B. E. (1989) *FEBS Lett.* 249, 243–247
4 Kemp, B. E., Bylund, D. B., Huang, T. S. and Krebs, E. G. (1975) *Proc. Natl Acad. Sci. USA* 72, 3448–3452
5 Kemp, B. E. (ed.) (1990) *Peptides and Protein Phosphorylation*, pp. 1–314, Uniscience CRC Press
6 Chan, K. F. J., Hurst, M. O. and Graves, D. J. (1982) *J. Biol. Chem.* 257, 3655–3659
7 Kemp, B. E. and Pearson, R. B. (1985) *J. Biol. Chem.* 260, 3355–3359
8 Pearson, R. B., Woodgett, J. R., Cohen, P. and Kemp, B. E. (1985) *J. Biol. Chem.* 260, 14471–14476
9 Ribadeau-Dumas, B., Grosclaude, F. and Mercier, J. C. (1970) *Eur. J. Biochem.* 14, 451–459
10 Brzeska, H., Lynch, T. J., Martin, B. and Korn, E. D. (1989) *J. Biol. Chem.* 264, 19340–19348
11 Czernik, A. J., Pang, D. T. and Greengard, P. (1987) *Proc. Natl Acad. Sci. USA* 84, 7518–7522
12 Nairn, A. C. and Palfrey, H. C. (1987) *J. Biol. Chem.* 262, 17299–17303
13 Graff, J. M., Stumpo, D. J. and Blackshear, P. J. (1989) *J. Biol. Chem.* 264, 11912–11919
14 Erikson, E. and Maller, J. L. (1988) *Second Messengers and Phosphoproteins* 12, 135–143
15 Colthurst, D. R., Campbell, D. G. and Proud, C. G. (1987) *Eur. J. Biochem.* 166, 357–363
16 Lees-Miller, S. P. and Anderson, C. W. (1989) *J. Biol. Chem.* 264, 17275–17280
17 Yeaman, S. J. (1990) *Biochim. Biophys. Acta* 1052, 128–131
18 Langan, T. A. (1978) *Methods Cell Biol.* 19, 143–152
19 Vulliet, P. R., Hall, F. L., Mitchell, J. P. and Hardie, D. G. (1989) *J. Biol. Chem.* 264, 16292–16298
20 Countaway, J. L., Northwood, I. C. and Davis, R. J. (1989) *J. Biol. Chem.* 264, 10828–10835
21 Fiol, C. J., Mahrenholz, A. M., Wang, Y., Roeske, R. W. and Roach, P. J. (1987) *J. Biol. Chem.* 262, 14042–14048
22 Hardie, D. G., Carling, D. and Sim, A. T. R. (1989) *Trends Biochem. Sci.* 14, 20–23
23 Edelman, A. M., Blumenthal, D. K. and Krebs, E. G. (1987) *Annu. Rev. Biochem.* 56, 567–613
24 Paxton, R., Kuntz, M. and Harris, R. A. (1986) *Arch. Biochem. Biophys.* 244, 187–201
25 Kudlicki, W., Wettenhall, R. E. H., Kemp, B. E., Szyszka, R., Kramer, G. and Hardesty, B. (1987) *FEBS Lett.* 215, 16–20
26 Rychlik, W., Russ, M. A. and Rhoads, R. E. (1987) *J. Biol. Chem.* 262, 10434–10437
27 Thorsness, P. E. and Koshland, D. E. (1987) *J. Biol. Chem.* 262, 10422–10425
28 Hausdorff, W. P., Bouvier, M., O'Dowd, B. F., Irons, G. P., Caron, M. G. and Lefkowitz, R. T. (1989) *J. Biol. Chem.* 264, 12657–12665
29 Palczewski, K., Arendt, A., McDowell, J. H. and Hargrave, P. A. (1989) *Biochemistry* 28, 8764–8770
30 Watson, M. H., Taneja, A. K., Hodges, R. S. and Mak, A. S. (1988) *Biochemistry* 27, 4506–4512

This article is the second in our series on 'conserved protein sequence motifs involved in recognition and targeting'. The series commenced in August 1990.

Int. J. Biochem. Vol. 18, No. 6, pp. 497–504, 1986
Printed in Great Britain.

0020-711X/86 $3.00 + 0.00

EXHIBIT 4

MINIREVIEW

MOLECULAR BASIS FOR SUBSTRATE SPECIFICITY OF PROTEIN KINASES AND PHOSPHATASES

JOHN W. SPARKS and DAVID L. BRAUTIGAN*
Section of Biochemistry, Brown University, Providence, RI 02912, U.S.A.

(*Received* 5 *November* 1985)

I. SUMMARY

Regulation of various metabolic processes occurs by the phosphorylation/dephosphorylation of enzymes. Both the protein kinases that catalyze the phosphorylations and the protein phosphatases that catalyze the dephosphorylations display relatively broad specificity, reacting with a number of distinct sites in target enzymes. In this way changes in the activity of a particular kinase or phosphatase can cause coordinated and pleiotropic responses. However, the kinases and phosphatases do not exhibit a one-to-one correspondence in their reactions. Residues at different positions may be phosphorylated by a single kinase, yet dephosphorylated by different individual phosphatases. Conversely, sites which are substrates for different individual kinases may be dephosphorylated by a single phosphatase. In exploring the molecular basis for these differences this article shows that whereas kinases react with specific primary structures that oftentimes appear as beta bends, the phosphatases recognize higher order structure, less strictly ruled by amino acid sequence surrounding the phosphorylated site. The differences, seen in the ability of these enzymes to utilize synthetic peptide substrates, might be rationalized in terms of function. Kinases need protruding segments of structure that can be enwrapped to exclude water, thereby minimizing ATP hydrolysis and enhancing phosphotransferase activity. On the other hand phosphatases are hydrolytic enzymes that may operate especially well on protein interfaces. Hydrolytic action often measured with *p*-nitrophenylphosphate is not necessarily indicative of a protein phosphatase and consideration of the mechanism reveals why this substrate can be misleading. Recognition of protein substrates by different, but overlapping, structural features separates the protein kinases and phosphatases and allows for precise control of the extent of phosphorylation of each site in target proteins by distinct pairs of interconverter enzymes.

II. SUBSTRATE SPECIFICITY OF PROTEIN KINASES

(a) *Protein Ser/Thr kinases*

During the past decade reversible phosphorylation of proteins has been recognized as an important mechanism by which many cellular metabolic activities are regulated. Studies have focused on the properties of two types of protein kinases; those which phosphorylate serine or threonine residues and those which phosphorylate tyrosine residues (see reviews Krebs and Beavo, 1979; Hunter and Sefton, 1982). Protein Ser/Thr kinases can be further classified into 3 subgroups according to the allosteric modifiers which regulate their activity. Of these groups, cAMP dependent protein kinase was the first to be classified and subsequently, cGMP and Ca^{2+}/calmodulin dependent protein kinases have also been characterized (Krebs and Beavo, 1979).

Initially, investigators were puzzled by the apparent lack of specificity of the protein kinases. For instance, the cAMP dependent protein kinase phosphorylated a variety of substrates including glycogen synthase, histone H1, the troponin inhibitory subunit and both the alpha and beta subunits of phosphorylase kinase (Soderling and Park, 1974). These proteins have quite different amino acid sequences and presumably unique tertiary structures. The principal features of the substrates responsible for determining kinase specificity remained an enigma until an important clue was provided by the observation that native chicken lysosyme, a poor substrate for the c-AMP dependent protein kinase, was rapidly phosphorylated subsequent to denaturation (Bylund *et al.*, 1975). Similarly, treatment of pyruvate kinase with strong base resulted in dramatic increase in the rate of its phosphorylation (Humble *et al.*, 1975). Thus, most investigators surmised that although there were primary structural determinants that would bind to the protein kinase, these were masked by other segments of the protein in the tertiary structure.

The importance of the amino acid sequence in determining substrate specificity of protein kinases has been demonstrated by a variety of approaches. Experiments using genetic variants of casein and peptide analogues of portions of the sequences surrounding phosphorylated residues of natural substrates have proven to be most illustrative. For instance, substitution of arginine for serine residue 122 in a variant of beta casein produced a 70-fold increase in the rate of phosphorylation at serine 124 (Kemp *et al.*, 1975). Thus, a single amino acid substitution can have a dramatic impact on specificity.

Use of synthetic peptides has been of great importance in identifying which residues are crucial for

*Author to whom correspondence should be addressed.

B.C. 18:6—A

substrate recognition. For example, the peptide Leu–Arg–Arg–Ala–Ser–Leu–Gly, an analogue of the sequence surrounding the phosphorylation site in pyruvate kinase served as a substrate for c-AMP dependent protein kinase with an apparent K_m of 16 mM (Kemp *et al.*, 1977). However, substitution of valine for either of the arginine residues, resulted in over a 100-fold increase in K_m (Kemp *et al.*, 1977). These and other results have shown that arginine residues preceding the site of phosphorylation are important specificity determinants for protein Ser/Thr kinases.

Peptides corresponding to the phosphorylation site in phosphorylase *b* have also been used to investigate the role basic and hydrophobic residues play in determining the specificity of phosphorylase *b* kinase, a calcium dependent enzyme. Substitution of valine for the lysine three residues amino terminal to the phosphorylated serine resulted in a five-fold decrease in V_{max}. Thus, in some cases, positively charged residues other than arginine function as key specificity determinants (Tessmer *et al.*, 1977). When hydrophobic residues adjacent to the phosphorylated serine were replaced by glycine (Tessmer *et al.*, 1977) or glutamic acid (Viriya and Graves, 1979) there was a two-fold increase in the apparent K_m values. Thus, hydrophobic residues adjacent to the phosphorylated serine are also important for the correct alignment of enzyme and substrate. Perhaps these hydrophobic residues facilitate the transfer of phosphate from ATP to protein by excluding water from the active site.

Does higher order structure of the substrate play a role in its recognition by kinases? Several studies indicate that regions of secondary structure of the substrate may act as important specificity determinants. Of thirty phosphorylated sites in proteins examined, twenty four of them were located in regions suspected to be beta turns (Small *et al.*, 1977). Thus, it was suggested that beta turns serve as important recognition sites for kinases. A subsequent study by Graves *et al.* (1978) supports this contention. When the peptide Leu–Ser–Tyr–Arg–Gly–Tyr–Ser–Leu was incubated with phosphorylase kinase or cAMP dependent kinase, serine 2 was phosphorylated by the former enzyme whereas serine 7 was phosphorylated by the latter one. This peptide contains the sequence Tyr–Arg–Gly–Tyr which forms a beta turn adjacent to the phosphorylated serine in heat denatured lysosyme. Graves *et al.* suggest that each kinase recognizes and binds to different sides of the beta turn which is presumably formed by the octapeptide.

In summary, the primary structure of the protein substrate plays a vital role in determining the specificity of these protein Ser/Thr kinases. Specifically, phosphorylation occurs on residues that are preceded by two or three basic amino acid sidechains. In many cases the phosphorylated residue is immediately surrounded by hydrophobic residues which presumably enhance phosphorylation by excluding water from the active site and promote correct alignment of substrates. It appears that beta turns of protein substrates, and thus secondary structure, plays an important role in recognition of the phosphorylatable residue by the kinases.

(b) *Protein Tyr kinases*

Over the last five years a second class of kinases specific for tyrosine residues has been discovered. Initially, phosphorylation of tyrosine was an activity associated with retroviral transforming proteins (see review Hunter and Sefton, 1982). More recently investigations have revealed that receptors for mitogenic hormones such as EGF, PDGF, and insulin phosphorylate themselves or exogenously added proteins on tyrosine residues upon binding their respective ligands (Ushiro and Cohen, 1980; Frackelton *et al.*, 1984; reviews by Cobb and Rosen, 1984; Heldin and Westermark, 1984). Preparations of membranes containing these kinases have been employed to phosphorylate tyrosine residues in a wide variety of proteins and peptides including myosin light chains, bovine serum albumin, casein, histone and Val 5 angiotensin (Gallis *et al.*, 1983; Shriner and Brautigan, 1984; Sparks and Brautigan, 1985; Wong and Goldberg, 1983; Foulkes *et al.*, 1983).

The reactivity of the protein Tyr kinases with a wide range of proteins is reminiscent of the situation with the protein Ser/Thr kinases. Again, primary structure of the substrate plays an essential role in determining the specificity of the protein Tyr kinases. Analysis of the sequence surrounding the phosphorylated tyrosine in $pp60^{src}$, the transforming protein produced by Rous Sarcoma Virus, revealed that several acidic sidechains precede the phosphorylated residue. In fact, comparison of $pp60^{src}$ with p90, another viral oncogene product, reveals a striking homology in the sequence surrounding the phosphorylated residue (Hunter, 1982). Thus, in contrast to the protein Ser/Thr kinases, protein Tyr kinases phosphorylate residues preceded by acidic sidechains.

Studies employing peptide analogues of the sequence surrounding the phosphorylation site in $pp60^{src}$ have confirmed the importance of acidic sidechains as specificity determinants for tyrosine kinases. For example, the peptide Lys–Leu–Ile–Glu–Asp–Asn–Glu–Tyr–Thr–Ala–Arg served as a substrate for the transforming protein of Rous Sarcoma Virus with an apparent K_m of 5 mM. Deletion of Glu at position 4 results in an 8-fold increase in the apparent K_m value (Hunter, 1982).

Recently it was observed that when myosin light chain 20 ($M_r = 20{,}000$) is incubated with the EGF/receptor kinase, Tyr 142 and Tyr 155 are phosphorylated with a stoichiometry near 1 mole of ^{32}P per mole of protein. However, residue 142 was phosphorylated three times faster (Gallis *et al.*, 1983). Inspection of the sequence surrounding each tyrosine reveals that several acidic residues are located 2–3 positions amino terminal to tyrosine 142 whereas there are no acidic sidechains in the sequence surrounding tyrosine 155 (Maita *et al.*, 1981).

What role does tertiary structure of the substrate play in determining protein Tyr kinase specificity? An observation in our laboratory indicates that it may restrict the kinase from potential sites, just as was the case for the c-AMP dependent kinase. We found that when reduced and alkylated serum albumin was incubated with the EGF/receptor kinase and [γ-^{32}P-ATP], nearly all of the radioactivity was incor-

porated into one tryptic peptide fragment, as revealed by reverse-phase HPLC mapping. However, when the albumin was digested with trypsin prior to phosphorylation by the EGF/receptor kinase, more than six peptides contained substantial amounts of radioactivity. Apparently portions of the tertiary structure mask specificity determinants contained in the amino acid sequence of the substrate, so the kinase only reacts with sites fully exposed on the surface.

Since the *in vivo* substrates for protein Tyr kinases continue to elude identification, most researchers have used artificial substrates *in vitro*, and the extent of phosphorylation is usually quite low. Thus, caution must be exercised when interpreting these results. However, it is intriguing that protein Ser/Thr kinases are specific for residues preceded by basic sidechains whereas protein Tyr kinases display a preference for residues preceded by several acidic sidechains.

Although the amino acid sequence surrounding the phosphorylated residue is an important specificity determinant for both types of protein kinases, it must be emphasized that primary structure alone does not determine overall reactivity. In many cases the apparent K_m values for peptide analogues are 10- to 1000-fold higher than those for protein substrates with more extensive tertiary structure. For example, c-AMP dependent protein kinase phosphorylates troponin with an apparent K_m of 21 μM. However, the K_m for the phosphorylation of a peptide corresponding to the phosphorylation site in native troponin is 2320 μM, approximately 100-fold higher (Kemp, 1979). Some of this difference can be attributed to the limited number of conformations possible for a segment of polypeptide in an intact protein, compared to one in solution. Nonetheless, features of the substrate structure other than the sequence of amino acids immediately surrounding the phosphorylated sidechain play a role in determining the specificity and reactivity of protein kinases.

III. SUBSTRATE SPECIFICITY OF THE PROTEIN PHOSPHATASES

(a) *Protein Ser(P)/Thr(P) phosphatases*

In contrast to the protein kinases much less is known about protein phosphatases. One reason for the lack of concrete data is that until very recently most protein phosphatases have not been distinguished from one another by useful functional criteria and therefore have remained poorly characterized. Another factor which contributes to the complexity of the field is that protein phosphatase activity is assayed under a variety of conditions with different phosphoprotein substrates. In fact, the number of different assays for phosphatase activity has generated confusion in the field.

As with the protein kinases, protein phosphatases can be grouped according to their specificity. There are two classes of cellular protein phosphatases; those specific for Ser(P) and Thr(P) residues, and those specific for Tyr(P) residues. Acid and alkaline phosphatases which display an extremely broad specificity for phosphoesters, also display protein phosphatase activity.

Once the specificity of the protein Ser/Thr kinases was elucidated, researchers began to focus their attention on the protein Ser(P) phosphatases. Although cellular Ser(P) phosphatase activity was initially observed over 40 years ago (Cori and Greene, 1943) progress on characterizing the enzymes responsible for this activity has been unusually slow.

Ingebritsen and Cohen (1983) have adopted a system which is useful for classifying cellular protein Ser(P) phosphatases. According to their scheme there are two types. Type 1 phosphatases preferentially dephosphorylate the beta subunit of phosphorylase kinase and are inhibited by nanomolar concentrations of inhibitor 1 and inhibitor 2, two heat stable proteins. In contrast to the type 1 phosphatases, type 2 phosphatases preferentially dephosphorylate the alpha subunit of phosphorylase kinase and are insensitive to the action of inhibitors 1 and 2.

Initial studies demonstrated that protein Ser(P) phosphatase purified from rat liver displayed an optimal activity at pH 7.5 and was able to dephosphorylate phosphopyruvate kinase and phosphoprotamine with apparent K_m values of 27 and 34 μM, respectively (Titanji, 1977). In contrast to acid or alkaline phosphatases this same preparation was unable to dephosphorylate beta glycerophosphate or D/L phosphoserine (Titanji, 1977). Thus it appears as if protein Ser(P) phosphatases require the structure provided by a polypeptide substrate.

Subsequent studies with peptides corresponding to the phosphorylated site in pyruvate kinase revealed that the rat liver phosphatase was able to dephosphorylate phosphopeptides. However, the apparent K_m values for the peptides were 2- to 20-fold higher than those for the native protein, confirming that protein Ser(P) phosphatases prefer protein substrates. In addition, when glutamic acid was substituted for the carboxy terminal glutamine in the peptide analogues, the apparent K_m values increased dramatically. Thus, negatively charged sidechains presumably weaken the binding to the phosphatase (Titanji *et al.*, 1980).

More recently it has been demonstrated that protein Ser(P) phosphatases types 1 and 2 can dephosphorylate over twenty different proteins with serine and threonine residues phosphorylated by eight different kinases (Ingebritsen and Cohen, 1983). Although most of the substrates contain several basic residues preceding the phosphorylated sidechain, thereby making them suitable substrates for the kinases, it does not appear as if positively charged residues act alone as specificity determinants for the phosphatases. This notion is supported by the observation that protein Ser(P)/Thr(P) phosphatases dephosphorylate Ser(P) myosin light chains. Inspection of the sequence surrounding the phosphorylated serine revealed that there are no basic residues in close proximity to the phosphorylated sidechain. Another intriguing observation is that both types of phosphatases dephosphorylate glycogen synthase sites 3b and 3c, two sites containing no basic sidechains immediately preceding the phosphorylated residues. However, the dephosphorylation at these sites was one tenth that of the dephosphorylation of phosphorylase kinase (Ingebritsen and Cohen, 1983). These results indicate that primary structure may not be as crucial for determining specificity for phos-

phatases as it is for kinases. However, elements of primary structure may impair substrate reactivity. For example, relative to the alpha and beta subunits of phosphorylase kinase, glycogen synthase site 5 is a 100-fold less reactive with protein Ser(P)/Thr(P) phosphatase types 1 and 2. This site is unique in that the phosphorylated serine is immediately followed by 5 acidic residues (Cohen *et al.*, 1982; Ingebritsen and Cohen, 1983). This observation confirms earlier observations by Titanji *et al.* (1980) that acidic residues carboxy terminal to the phosphorylated sidechain may act as negative specificity determinants.

In summary, protein Ser(P)/Thr(P) phosphatases exhibit a broad specificity for a variety of protein substrates. They apparently recognize both acidic and basic proteins with diverse sequences. In contrast to the kinases, protein Ser(P) phosphatases do not display an absolute requirement for basic residues preceding the phosphorylated serine. Thus, phosphorylation/dephosphorylation at a site in a protein substrate might be catalyzed by a distinct kinase/phosphatase pair. Another site phosphorylated by the same kinase, because of its same primary structure, might be dephosphorylated exclusively by a different phosphatase (Fig. 1). This distinction may have contributed to the lack of understanding of phosphatase specificity and the failure to assign useful functional classifications to phosphatases in the past. It appears that elements of tertiary structure of the substrate play a critical role in recognition by the phosphatases. It is important to emphasize that while protein Ser(P)/Thr(P) phosphatases are reactive with both phosphoproteins and phosphopeptides, they have not been found to dephosphorylate small phosphorylated metabolites such as D/L phosphoserine and β-glycerophosphate.

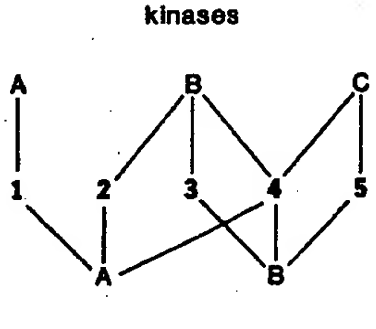


Fig. 1. Specificity of protein kinases and phosphatases. The reactivity of kinases A, B and C (top) and phosphatases A and B (bottom) with a number of individual protein substrate sites, 1 through 5, is indicated by the lines. This figure schematically shows how these enzymes have broad specificity, yet do not display a one-to-one correspondence in their reactivity. In this way changes in the activity of either a kinase or phosphatase would produce distinct metabolic effects. Presumably the molecular basis for this pattern is that kinases and phosphatases use different but overlapping structural features of the substrates for recognition. The examples used in this figure are kinase A, myosin light chain kinase; kinase B, cAMP-dependent protein kinase; kinase C, phosphorylase kinase; phosphatase A, type-2 and B, type-1 according to Ingebritsen and Cohen. Substrates are: 1, myosin light chains; 2, phosphorylase kinase alpha subunit; 3, phosphorylase kinase beta subunit; 4, glycogen synthase; 5, phosphorylase.

(b) *Protein Tyr(P) phosphatases*

Shortly after the discovery of protein tyrosine kinases, several laboratories including our own set out to purify protein Tyr(P) phosphatases. Recently, cytosolic protein Tyr(P) phosphatase activity from extracts of chicken brains was characterized. The major activity was eluted from a DEAE column at 145 mM NaCl, had an apparent molecular weight of 43,000, displayed optimal activity at pH 7.0 and was inhibited by micromolar concentrations of Zn^{2+}. Relative to bovine intestinal alkaline phosphatase, chicken brain protein Tyr(P) phosphatase was nearly 3000 times more reactive with phosphocasein than with *para*-nitrophenylphosphate (Foulkes *et al.*, 1983).

Using Zn^{2+} agarose affinity chromatography followed by DEAE-Sepharose chromatography, two types of protein Tyr(P) phosphatases were purified to near homogeneity from rabbit kidney. Protein Tyr(P) phosphatase peak I displayed optimal activity at pH 7.0 and required mercaptans for activity. In contrast the peak II phosphatase displayed optimal activity at pH 5.0 and did not require mercaptans for activity (Shriner and Brautigan, 1984). Both enzymes, however, displayed parallel substrate specificity. For instance protein Tyr(P) phosphatases were unreactive with Ser(P) proteins but dephosphorylated Tyr(P) albumin and Tyr(P) casein with apparent K_m values ranging from 2.2 to 15 μM (Sparks and Brautigan, 1985). In addition, both phosphatases were able to dephosphorylate Tyr(P) EGF receptors in preparations of A431 membranes. Surprisingly, Tyr(P) histone, a basic protein, was not dephosphorylated by either enzyme (Sparks and Brautigan, 1985). Furthermore, protein Tyr(P) phosphatases were not able to dephosphorylate Tyr(P) angiotensin and with Tyr(P) RR-src, a peptide analogue of the phosphorylated site in the transforming protein $pp60^{src}$, less than 12% of the phosphorous was removed from the peptide by either enzyme.

At this point, it is interesting to note that all the Tyr(P) proteins which serve as substrates for these phosphatases have acidic isoelectric points. Thus, it appears as if protein Tyr(P) phosphatases recognize acidic substrates and require some defined tertiary structure.

(c) *Acid and alkaline phosphatases*

The final group of phosphatases that must be discussed because of their reactivity with protein substrates were initially characterized in the 1930's. In spleen extracts, investigators observed a phosphatase activity which could hydrolyze β-glycerophosphate as well as several alkyl phosphates. This phosphatase activity displayed optimal reactivity at both pH 4.8 and 9.0 and the enzymes responsible for the observed hydrolysis were classified as acid and alkaline phosphatase, respectively (Davies, 1934).

Since then a great deal has been learned about alkaline phosphatase present in a variety of mammalian tissues. There are two classes of mammalian

alkaline phosphatases; those which require Zn^{2+} and are stimulated by Mg^{2+} and those which require Zn^{2+} and are unaffected by other divalent metals (Cathala and Brunel, 1974).

Both classes of alkaline phosphatases have similar specificities. They display an optimal activity between pH 8 and 10 and are able to dephosphorylate a variety of small molecules including *p*-nitrophenylphosphate. For instance, it has been demonstrated that bovine kidney alkaline phosphatase dephosphorylates α-naphthylphosphate, glucose 1-phosphate and β-glycerophosphate with apparent K_m values of 1.8, 128 and 68 μM, respectively (Cathala and Brunel, 1974). These values contrast the millimolar K_m values observed with peptide substrates for protein phosphatases.

More recently, investigators have revealed that mammalian alkaline phosphatases are able to dephosphorylate a variety of Ser(P) and Tyr(P) proteins. One study revealed that placental alkaline phosphatase dephosphorylates Ser(P) histone and Ser(P) protamine with apparent K_m values of 10 and 5 μM, respectively (Huang *et al.*, 1976). A subsequent study revealed that both calf intestine and bovine liver alkaline phosphatase preferentially dephosphorylate Tyr(P) histones relative to Ser(P) histones (Swarup *et al.*, 1981).

The three classes of cellular phosphatases can be distinguished by the effects of compounds that selectively inhibit enzyme activity. Thus, specific inhibition may be a useful criterion for enzyme identification. Several studies have shown that protein Ser(P)/Thr(P) phosphatases are inhibited by millimolar concentrations of ATP and GTP (Krebs *et al.*, 1976; Hsiao, 1978), inorganic pyrophosphate (Khandelwal and Kamani, 1980; Yan and Graves, 1982) and fluoride (Khatra and Soderling, 1978). In contrast to the protein Ser(P)/Thr(P) phosphatases, the protein Tyr(P) phosphatases and the alkaline phosphatases are unaffected by fluoride and only weakly inhibited by inorganic pyrophosphate. Instead protein Tyr(P) phosphatases are potently inhibited by micromolar concentrations of orthovanadate (Swarup *et al.*, 1982) or micromolar concentrations of Zn^{2+} (Brautigan *et al.*, 1981), whereas alkaline phosphatases are inhibited by orthovanadate but require Zn^{2+} for catalytic activity (Lopez *et al.*, 1976; Fernley, 1971). In fact, unlike the protein Tyr(P) phosphatases which show optimum activity with EDTA, the alkaline phosphatases are inhibited by divalent cation chelators such as EDTA (Fernley, 1971).

In summary, the three types of phosphatases exhibit different patterns of inhibition, so inhibitors such as Zn^{2+}, vanadate, fluoride and EDTA may be useful for classifying a phosphatase under investigation. More importantly, inhibition of different types of phosphatases by different ions may reflect fundamental differences in the structure of their respective active sites.

(d) *Phosphatase activity and specificity: the special case of* p-*nitrophenyl phosphate*

As interest in phosphatases has intensified, many investigators have adopted *para*-nitrophenyl phosphate (*p*NPP) as substrate in assays for phosphatase activity. Upon hydrolysis of *p*NPP, the yellow *p*-nitrophenolate ion is produced. The concentration of *p*-nitrophenolate ions is usually monitored by stopping the reaction with Na_2CO_3 and measuring the absorbance of the solution at 410 nm. This method has been adopted widely as a rapid and sensitive colorometric assay for detecting phosphatase activity. However, divalent metals are often added to the assay mixture in millimolar concentrations (Li, 1979; Foulkes *et al.*, 1983). Analysis of phosphate ester hydrolysis reveals that a variety of metals including biologically important divalent cations such as Mg^{2+} and Zn^{2+} enhance the non-enzymatic hydrolysis of phosphate esters (Steffins *et al.*, 1973). In assays for phosphatase activity, where both a divalent metal and *p*NPP are present in equimolar concentrations, the effective substrate may be a metal–*p*NPP complex which is particularly susceptible to hydrolysis. Under such conditions the use of *p*NPPase activity as the sole criterion for classifying a cellular enzyme as a phosphatase may lead to erroneous conclusions. A brief discussion of phosphate ester hydrolysis is illustrative and useful in addressing the following two questions concerning phosphatase specificity: (1) Is *para*-nitrophenyl phosphate hydrolysis a useful criterion for the unambiguous classification of cellular phosphatases? (2) Is *p*NPP hydrolysis an inherent activity of all protein phosphatases?

Early studies of phosphate esters indicated that during hydrolysis cleavage occurs at the P–O bond and not the C–O bond. This fact is supported by two lines of evidence. First, ^{18}O exchange experiments with phosphate monoesters revealed that ^{18}O is incorporated into phosphate but not into the oxygen of the alcohol produced during hydrolysis. Second, there is retention of stereochemical configuration of the alcohol produced during hydrolysis (Butcher and Westheimer, 1955). These results are consistent with a mechanism where the phosphorous undergoes nucleophilic attack by water during hydrolysis. It is important to note that the phosphate–oxygen bond is also cleaved during the hydrolysis of β-glycerophosphate by alkaline phosphatases (Stein and Koshland, 1952). Alkaline phosphatases display a stringent requirement for a tightly bound Zn^{2+} ion at their active sites. Presumably, the metal facilitates the hydrolysis of phosphate esters by polarizing the phosphorous atom, thereby enhancing its susceptibility to nucleophilic attack.

Another intriguing property of phosphate monoesters is that they are rapidly hydrolyzed under acidic conditions but are relatively stable under mild alkaline conditions. At pH 8–9 the phosphate dianion is the most prevalent species in solution, thus one expects that hydrolysis is retarded by simple electrostatic repulsion between the phosphate and hydroxide ions (Westheimer and Schookhoff, 1940). However, studies with lanthanum (La) hydroxide gels (Butcher and Westheimer, 1955) have demonstrated that La enhances hydrolysis of phosphate esters at pH 8.5, conditions where the dianionic form predominates. Presumably, La is coordinated with the negatively charged oxygens as well as the alkoxide ion of the leaving group in a transient intermediate, thereby promoting hydrolysis.

Likewise, with phosphate diesters (Steffins *et al.*, 1973) it was observed that biologically important

metals enhanced intramolecular hydrolysis. Specifically it was observed that Zn^{2+} and Mg^{2+} enhanced by 304-fold and 24-fold the rate of hydrolysis of $^{-}OOC-CH(CH_3)(OPO_3C_6H_5)$. This enhancement was attributed to the stabilization of the presumed pentacovalent intermediate by the divalent cation. The phosphodiester was envisioned as having a trigonal bipyramid configuration with the nucleophile and leaving group occupying coplanar apical positions in the intermediate produced during hydrolysis. An elegant study with *para*-nitrophenylphosphate confirmed that metals do indeed promote intermolecular hydrolysis of phosphate monoesters. Chelated *p*-nitrophenylphosphatobis (trimethylenediamine) cobalt III was hydrolyzed 10^5-fold more rapidly than free *p*NPP at pH 10.2 (Anderson *et al.*, 1977). This rate enhancement was attributed to strain relief in the pentacovalent intermediate generated during hydrolysis. Additional rate enhancement may also arise as a consequence of charge neutralization. Thus, hydroxide ions or other nucleophiles in solution will more readily attack a phosphate ester coordinated to a positively charged metal ion than a phosphate dianion.

These kinetic studies of phosphate ester hydrolysis provide evidence that metal ions enhance catalysis presumably by stabilizing a pentacovalent intermediate. These results have profound implications for phosphatase assays which include divalent metals in an alkaline reaction mixture.

A few additional characteristics of pNPP merit further discussion. Relative to the alkyl phosphates glucose-6-phosphate and β-glycerophosphate, two compounds which are readily hydrolyzed by alkaline phosphatases (Cathala and Brunel, 1975), *p*NPP releases approximately 2–3 kcal/mol more free energy upon hydrolysis. The primary factor contributing to the release of more free energy is that a nitrophenolate anion is more stable than an alkoxyl anion. This fact is reflected in the relative pK_a values for PNP and aliphatic alcohols which are 8.5 and *c.* 15.0, respectively. The unusual stability of the *p*-nitrophenolate ion enhances the enzymatic hydrolysis of *p*NPP, for during hydrolysis the nitrophenolate ion does not have to be protonated by a conjugate acid at the active site of the enzyme. In contrast, alkyl phosphates such as protein serine phosphate or glucose 6-phosphate display a strict requirement for enzymatic protonation of the alcohol leaving group, an important aspect of the mechanism of hydrolysis of phosphate esters by phosphatases. Enzymes which are unreactive with alkylphosphates because of their inability to protonate the alkoxide ion leaving group nonetheless may be able to hydrolyze *p*NPP.

Taken together, the elimination of the need to protonate the leaving group and the acceleration of hydrolysis by metal ions, suggest that Mg·*p*NPP complexes are hydrolyzed rather easily and may be susceptible to enzymes that are unrelated to phosphatases. Therefore one must exercise caution when interpreting the results of an assay which relies solely on the hydrolysis of *p*NPP to detect phosphatase activity. In fact, several investigators have observed that *p*NPP can be hydrolyzed by enzymes which are distinct from phosphatases in both structure and function.

As an illustrative example, when *p*NPP was incubated in Tris buffer, pH 8.5 containing 10 mM *p*NPP and 10 mM $MgCl_2$, hexokinase hydrolyzed it with an apparent velocity ranging from 3.40 to 5.14 nmol/min/mg. However, in a parallel experiment without $MgCl_2$ no hydrolysis was observed. This intriguing observation strongly suggests that equimolar concentrations of Mg^{2+} ions and *p*NPP lead to the formation of a complex which is easily hydrolyzed by certain proteins in solution. Hexokinase also can hydrolyze *p*NPP at pH 7.0. In HEPES buffer containing equimolar concentrations of Mg^{2+} and *p*NPP, *p*NPP is hydrolyzed with an apparent velocity of 6.6 nmol/min/mg. As in the experiment described above, the hydrolysis of *p*NPP by hexokinase displayed a stringent requirement for Mg^{2+} ions (unpublished observations). It is important to note that the ability of hexokinase to behave as a phosphatase under the conditions described above strongly supports our contention that millimolar concentrations of Mg^{2+} enhance the hydrolysis of *p*NPP by forming a Mg·*p*NPP complex. This complex is extremely labile and may be hydrolyzed by other enzymes in addition to hexokinase that are unrelated to the phosphatases. Several years ago, Koester *et al.* (1980) demonstrated that carbonic anhydrase III hydrolyzed *p*NPP with an apparent K_m value of 4.17 mM. This observation is surprising since Koester *et al.* excluded metal ions from their assay mixtures.

As to the second question, since Mg–*p*NPP complexes appear to be so labile and the mechanistic requirements for *p*NPP hydrolysis so lax one might expect all phosphatases to readily hydrolyze *p*NPP. However, some but not all phosphatases are highly reactive with *p*NPP. Many investigators have found that alkaline phosphatases, enzymes containing tightly bound Zn^{2+}, rapidly hydrolyze *p*NPP (Fernley, 1971). Several recent studies have revealed that Tyr(P) phosphatases are also reactive with *p*NPP. For instance, three types of Tyr(P) phosphatase activity isolated from chicken brain extracts hydrolyze *p*NPP (Foulkes *et al.*, 1983). In our laboratory we have observed that protein Tyr(P) phosphatases purified from rabbit skeletal muscle display a *p*NPPase activity. This activity is divalent cation independent the assay mixture contains 10 mM *p*NPP and 2 mM EDTA but no Mg^{2+} (Shriner and Brautigan, 1984). The reactivity of protein Tyr(P) phosphatases with *p*NPP is not surprising since the structure of *p*NPP resembles that of phosphotyrosine. In contrast, the protein Ser(P) phosphatases have much lower reactivity with *p*NPP (Brautigan *et al.*, 1982).

In summary, *p*NPP is much more susceptible to enzymatic hydrolysis than alkyl phosphate monoesters. It should not be used alone to classify an enzyme as a phosphatase, especially a protein phosphatase, because other unrelated proteins can show substantial levels of activity that are potentially misleading. Although all protein phosphatases might display some level of inherent *p*NPPase activity, the amount of activity will be considerably different for different members of this family and the results strongly dependent on the conditions employed, in particular the concentration of divalent cations.

Acknowledgement—This work was supported by NIH Grant AM31374.

REFERENCES

Anderson B., Milburn R., Harrowfield J., Robertson G. and Sargenson A. (1977) Cobalt(III)-promoted hydrolysis of a phosphate ester. *J. Am. Chem. Soc.* **99**, 2652–2661.

Brautigan D., Bornstein P. and Gallis B. (1981) Phosphotyrosyl-protein phosphatase. *J. biol. Chem.* **256**, 6519–6522.

Brautigan D., Khatra B., Soderling T. and Fischer E. (1982) Dephosphorylation of skeletal muscle phosphorylase, glycogen synthase, and phosphorylase kinase beta-subunit by a Mn^{2+}-activated protein phosphatase. *Archs Biochem. Biophys.* **219**, 228–235.

Butcher W. and Westheimer F. (1955) The lanthanum hydroxide gel promoted hydrolysis of phosphate esters. *J. Am. Chem. Soc.* **77**, 2420–2424.

Bylund D. and Krebs E. (1975) Effect of denaturation on the susceptibility of proteins to enzymatic phosphorylation. *J. biol. Chem.* **250**, 6355–6361.

Cathala G. and Brunel C. (1975) Bovine kidney alkaline phosphatase. *J. biol. Chem.* **250**, 6046–6053.

Chan J., Hurst M. and Graves D. (1982) Phosphorylase kinase specificity. *J. biol. Chem.* **257**, 3655–3659.

Cobb M. and Rosen O. (1984) The insulin receptor and tyrosine protein kinase activity. *Biochim. biophys. Acta* **738**, 1–8.

Cohen P., Yellowlees D., Aitken A., Donella-Deana A., Hemmings B. and Parker P. (1982) Separation and characterization of glycogen synthase kinase 3, glycogen synthase kinase 4 and glycogen synthase kinase 5 from rabbit skeletal muscle. *Eur. J. Biochem.* **124**, 21–35.

Cori G. and Greene A. (1943) Crystalline muscle phosphorylase. *J. biol. Chem.* **151**, 31–38.

Davies D. (1934) The phosphatase activity of spleen extracts. *Biochem. J.* **28**, 529–536.

Fernley W. Mammalian alkaline phosphatases. In *The Enzymes* (Edited by Boyer P. D.), 3rd edn, Vol. IV, p. 417. Academic Press, New York.

Foulkes J., Erikson E. and Erikson R. (1983) Separation of multiple phosphotyrosyl- and phosphoseryl-protein phosphatases from chicken brain. *J. biol. Chem.* **258**, 431–438.

Frackelton A., Trenble P. and Williams L. (1984) Evidence for the platelet-derived growth factor-stimulated tyrosine phosphorylation of the platelet-derived growth factor receptor *in vivo*. *J. biol. Chem.* **259**, 7907–7915.

Gallis B., Edelman A., Casnellie J. and Krebs E. (1983) Epidermal growth factor stimulates tyrosine phosphorylation of the myosin regulatory light chain from smooth muscle. *J. biol. Chem.* **258**, 13089–13093.

Graves D., Uhing J., Janski A. and Viriya J. (1978) Use of a double-headed peptide substrate to study the specificity of cAMP-dependent protein kinase and phosphorylase kinase. *J. biol. Chem.* **253**, 8010–8012.

Heldin C. and Westermark B. (1984) Growth factors: Mechanism of action and relation to oncogenes. *Cell* **37**, 9–20.

Hsiao K., Sandberg A. and Li H. (1978) The role of ATP and divalent cations in the regulation of a cardiac phosphorylase phosphatase (phosphoprotein phosphatase) of $M_r = 35{,}000$. *J. biol. Chem.* **253**, 6901–6907.

Huang K., Robinson J. and Chou J. (1976) Phosphoprotein-phosphatase activity associated with human placental alkaline phosphatase. *Biochem. biophys. Res. Commun.* **70**, 186–192.

Humble E., Berglund L., Titanji V., Ljungstrom O., Edlund B., Zetterquist O. and Engstrom L. (1975) Non-dependence on native structure of pig liver pyruvate kinase when used as a substrate for cyclic 3',5'-AMP-stimulated protein kinase. *Biochem. biophys. Res. Commun.* **66**, 614–621.

Hunter T. (1982) Synthetic peptide substrates for a tyrosine protein kinase. *J. biol. Chem.* **257**, 4843–4848.

Hunter T. and Sefton B. (1982) Protein kinases and viral transformation. In *Molecular Action of Toxins and Viruses* (Edited by Cohen P. and van Heyningen S.), pp. 333–366. Elsevier/North-Holland Biomedical Press, Amsterdam.

Ingebritsen T. and Cohen P. (1983) The protein phosphatases involved in cellular regulation: I. Classification and substrate specificities. *Eur. J. Biochem.* **132**, 255–261.

Kemp B. (1979) Phosphorylation of synthetic peptide analogs of rabbit cardiac troponin inhibitory subunit by the cyclic AMP-dependent protein kinase. *J. biol. Chem.* **254**, 2638–2642.

Kemp B., Bylund D., Huang T. and Krebs E. (1975) Substrate specificity of the cyclic AMP-dependent protein kinase. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3448–3452.

Kemp B., Graves D., Benjamini E. and Krebs E. (1977) Role of multiple basic residues in determining the substrate specificity of cyclic AMP-dependent protein kinase. *J. biol. Chem.* **252**, 4888–4894.

Khandelwal R. and Kamani S. (1980) Studies on inactivation and reactivation of homogeneous rabbit liver phosphoprotein phosphatases by inorganic pyrophosphate and divalent cations. *Biochim. biophys. Acta* **613**, 95–105.

Khatra B. and Soderling T. (1978) Reversible inhibition of skeletal muscle phosphoprotein phosphatase by ATP, phosphate and fluoride. *Biochem. biophys. Res. Commun.* **85**, 647–654.

Koester M., Pullan L. and Noltmann E. (1981) The *p*-nitrophenyl phosphatase activity of muscle carbonic anhydrase. *Archs Biochem. Biophys.* **211**, 632–642.

Krebs E. and Beavo J. (1979) Phosphorylation-dephosphorylation of enzymes. *A. Rev. Biochem.* **48**, 923–959.

Krebs E., Kandelwal R. and Vandenheede J. (1976) Purification properties and substrate specificities of phosphoprotein phosphatases from rabbit liver. *J. biol. Chem.* **251**, 4850–4858.

Li H. (1979) Purification and properties of a phosphorylase (phosphoprotein) phosphatase associated with an alkaline phosphatase of M_r 35,000 from bovine adrenal cortex. *Eur. J. Biochem.* **102**, 363–374.

Lopez V., Stevens T. and Lindquist R. (1976) Vanadium ion inhibition of alkaline phosphatase-catalyzed phosphate ester hydrolysis. *Archs Biochem. Biophys.* **175**, 31–38.

Maita T., Chen J. and Matsuda G. (1981) Amino-acid sequence of the 20,000-molecular-weight light chain of chicken gizzard-muscle myosin. *Eur. J. Biochem.* **117**, 417–424.

Shriner C. and Brautigan D. (1984) Cytosolic protein phosphotyrosine phosphatases from rabbit kidney. *J. biol. Chem.* **259**, 11383–11390.

Small D., Chou P. and Fasman G. (1977) Occurrence of phosphorylated residues in predicted beta-twins: Implications for beta twin participation in control mechanisms. *Biochem. biophys. Res. Commun.* **79**, 341–346.

Soderling T. R. and Park C. R. (1974) Recent advances in glycogen metabolism. In *Advances in Cyclic Nucleotide Research* (Edited by Greengard P. and Robinson G. A.), Vol. 4, pp. 283–333. Raven Press, New York.

Sparks J. and Brautigan D. (1985) Specificity of protein phosphotyrosine phosphatases. *J. biol. Chem.* **260**, 2042–2045.

Steffens J., Sampson E., Siewers I. and Benkovic S. (1973) Effect of divalent metal ions on the intramolecular nucleophilic catalysis of phosphate diester hydrolysis. *J. Am. Chem. Soc.* **95**, 936–938.

Stein S. and Koshland D. (1952) Mechanism of action of alkaline phosphatase. *Arch. Biochem. Biophys.* **39**, 229–230.

Swarup G., Cohen S. and Garbers D. (1981) Selective

dephosphorylation of proteins containing phosphotyrosine by alkaline phosphatases. *J. biol. Chem.* **256**, 8197–8201.

Swarup G., Cohen S. and Garbers D. (1982) Inhibition of membrane phosphotyrosyl-protein phosphatase activity by vanadate. *Biochem. biophys. Res. Commun.* **107**, 1104–1109.

Tessmer G., Skuster J., Tabatabai L. and Graves D. (1977) Studies on the specificity of phosphorylase kinase using peptide substrates. *J. biol. Chem.* **252**, 5666–5671.

Titanji V. (1977) Purification and properties of a phosphoprotein phosphatase from rat liver. *Biochim. biophys. Acta* **481**, 140–151.

Titanji V., Ragnarsson U., Humble E. and Zetterqvist O. (1980) Phosphopeptide substrates of a phosphoprotein phosphatase from rat liver. *J. biol. Chem.* **255**, 11339.

Ushiro H. and Cohen S. (1980) Identification of phosphotyrosine as a product of epidermal growth factor-activated protein kinase in A-431 cell membranes. *J. biol. Chem.* **252**, 8363–8365.

Viriya J. and Graves D. (1979) Phosphorylation of synthetic peptide analogs of the phosphorylatable site of phosphorylase *b* with phosphorylase kinase. *Biochem. biophys. Res. Commun.* **87**, 17–24.

Westheimer F. and Shookhoff M. (1940) The electrostatic influence of substituents on reaction rates. *J. Am. Chem. Soc.* **62**, 269.

Wong T. and Goldberg A. (1983) *In vitro* phosphorylation of angiotensin analogs by tyrosyl protein kinases. *J. biol. Chem.* **258**, 1022–1025.

Yan S. and Graves D. (1982) Inactivation and reactivation of phosphoprotein phosphatase. *Molec. Cell. Biochem.* **42**, 21–29.

EXHIBIT 5

10

Assay of protein kinases and phosphatases using specific peptide substrates

MARIA RUZZENE and LORENZO A. PINNA

1. Introduction

The successful and widespread use of peptide substrates for the assay of protein kinases is grounded on the observation that the great majority of these enzymes are able to recognize phosphoacceptor sites defined by local structural features within their target proteins. Although this local sequence specificity is not the only tool ensuring the selectivity of protein kinases (indeed, in some cases, notably protein–tyrosine kinases, it may not even be a major one) many members of the protein kinase family phosphorylate short peptides reproducing their natural phosphoacceptor sites. The advantage of synthetic peptides, compared with protein substrates, for kinase assays and specificity studies include:

- they can be readily obtained in large amounts
- they have a well-defined chemical composition
- they are readily amenable to chemical modifications aimed at rendering their phosphorylation more efficient, more specific, and easier to test

However, the kinetics of peptide phosphorylation compared with those of the intact protein, and the relevance of local structural features, are quite variable for different protein kinases. The majority of protein–serine/threonine kinases (PSKs) are markedly site specific, and consequently display high catalytic efficiency towards relatively short peptides, provided that the peptides include the required specificity determinants. This approach provides a tool for tailoring peptides that are specific for a given Ser/Thr kinase, as well as allowing a sensitive assay of its activity. The site specificity of most protein–tyrosine kinases (PTKs) is not as clear-cut as that of Ser/Thr kinases. This is probably because factors other than the sequence around the phosphoacceptor sites are more important for tyrosine kinases' specificity, especially

the presence of protein–protein recognition modules outside the catalytic domain, such as SH2, SH3, PTB, and PH domains (reviewed in ref. 1). This may frustrate attempts to develop *specific* peptide substrates for individual PTKs. However, it does not prevent the sensitive assay of their activity using peptide substrates, whose kinetic parameters are sometimes comparable to the best peptide substrates of PSKs. The only shortcoming in these cases would be that the peptide substrates of PTKs are, generally speaking, more promiscuous than those of PSKs. This does not rule out the possibility of discrimination between different PTKs using more than one peptide substrate. While it is highly unlikely that an individual peptide will be completely specific for a given PTK, the activity ratio towards two distinct peptides can still be drastically different for two different PTKs.

A small set of highly specific protein kinases display phosphorylation rates with peptide substrates that are negligible compared with those with the parent protein, and, consequently, peptide substrates are of little use in these cases. Some calmodulin-dependent protein kinases (e.g. calmodulin-dependent protein kinase IV), the eIF2α kinases, the G-protein-coupled receptor kinases, Raf, and the dual specificity MAP kinase kinases (MEKs) fall into this category. It is conceivable that, in the future, effective peptide substrates for these dedicated protein kinases could be developed, albeit possibly not related to the sequence of the natural protein substrates. Pertinent to this are observations that, on the one hand, the mutation of protein phosphoacceptor sites may improve their phosphorylation over that of the wild type, and, on the other hand, that crucial determinants established with peptide substrates may prove to be of lesser importance with the intact protein substrate. An example is provided by C-terminal Src protein kinase (Csk) a dedicated PTK very poorly active towards even large peptides reproducing its natural target, but which phosphorylates unrelated peptides with high efficiency (2). As pointed out in the next section, the peptide library approach (see Chapter 16) could prove especially helpful to select peptide substrates suitable for assaying protein kinases in this class, as well as protein kinases whose natural targets are still unknown.

2. The residue and sequence specificity of protein kinases

Efficient phosphorylation of amino acids by protein kinases requires their incorporation into a discrete peptide sequence. During the catalytic event, the residue undergoing phosphorylation will bind to the active site, whereas the surrounding residues will interact with elements of the kinase outside the *sensu stricto* catalytic site. These latter have been identified in some instances by site-directed mutagenesis (e.g. 3–6) or by crystal structure analysis (7, 8) Both the nature of the phosphorylatable amino acid, and of the residues

surrounding it, will therefore influence the catalytic parameters, giving rise to 'residue specificity' and 'sequence specificity', respectively (reviewed in ref. 1). From a practical standpoint, the residue specificity gives rise to just two main types of residue selection, corresponding to the two major classes of protein kinases, i.e. serine/threonine-specific (PSKs) or tyrosine-specific (PTKs). In most instances the border between PSKs and PTKs, which can be predicted on the basis of primary structure motifs, is well defined and can be empirically drawn using peptide substrates of the two sorts, exclusively phosphorylated by either PSKs or PTKs. In contrast, the features that underlie the preference for serine or threonine within the PSK class remain unclear. As a general rule, serine residues are preferred over threonine residues, in some cases to such an extent that it would be tempting to postulate a subclass of serine-specific protein kinases inactive on threonine (e. g. the Golgi casein kinase and $p90^{S6K}$). In contrast, other PSKs (e.g. AMP-activated protein kinase and Cdc2) tolerate Ser↔Thr substitutions quite well. These differences have sometimes been exploited to improve the selectivity of peptide substrates (e.g. see ref. 9).

The selectivity of most Ser/Thr kinases, however, relies mainly on *sequence specificity*, often expressed by a so-called 'consensus sequence'. The elements of the sequence that are especially required to ensure efficient phosphorylation are referred to as 'specificity determinants', whose nature (generally either basic, hydrophobic or acidic residues, or proline) and position depend on the kinase considered. The relevance of individual determinants can also vary, indicating that some are more dispensable than others. The indispensable residues are highlighted in the common representations of the consensus sequence, while additional features that are not strictly required (albeit that they may substantially improve the phosphorylation efficiency) are often neglected. Likewise, not enough emphasis is always placed on *negative determinants*, whose presence can compromise the phosphorylation of otherwise suitable sites, and which can be exploited to improve the selectivity of peptides, as they can prevent phosphorylation by kinases other than the desired one(s).

The consensus sequences of a number of representative protein kinases are shown in *Table 1*. Most of them have been drawn from the analysis of natural phosphoacceptor sites, in conjunction with kinetic studies with peptides reproducing these, either exactly or with suitable modifications. This approach proved especially rewarding with many PSKs, while with PTKs *sensu stricto* consensus sequences consistent with the structure of natural phosphoacceptor sites were established only in a few cases. It should be stressed, however, that even in the case of PSKs, consensus sequences do not represent an absolute rule: atypical sites, where the specificity determinants are located at positions different from those highlighted in the consensus, are often phosphorylated with an efficiency comparable to 'canonical' sites.

Table 1. Consensus sequences of some protein kinases[a]

Basophilic PSKs	Consensus sequences[b]
PKA	R–(R/K)–X–(**S/T**)–B
PKG	R–(R/K)–X–(**S/T**)–B
PKC	(R/K)–(R/K)–X–(**S/T**)–B–(R/K)–(R/K)
$p70^{S6K}$	(K/R)–X–R–X–X–(**S/T**)–B
$p90^{S6K}$/MAPKAPK1	X–X–(R/K)–X–R–X–X–**S**–X–X
	R–R–R–X–**S**–X–X
MAPKAPK2	X–X–B–X–R–X–X–**S**–X–X
CaMK I	B–X–R–X–X–(**S/T**)–X–X–X–B
CaMK II	B–X–(R/K)–X–X–(**S/T**)–X–X
Phosphorylase kinase	K–R–K–Q–I–**S**–V–R
AMPK	B–(X,R/K/H)–X–X–(**S/T**)–X–X–X–B
HSV-PK and PRV-PK	R–R–R–R–X–(**S/T**)–X
Proline–directed PSKs	
ERK (MAP kinases)	P–X–(**S/T**)–P–P
Cdc2 (and other cyclin-dependent PKs)	X–(**S/T**)–P–X–(K/R)
Acidophilic PSKs	
CK2	X–(**S/T**)–X–(E/D/Sp/Yp)–X
G-CK	X–**S**–X–(E/Sp)–X
CK1	(Sp/Tp)–X–X–(**S/T**)–B
	$(D/E)_n$–X–X–(**S/T**)–X–X ($n \geq 4$)
GSK-3	(**S/T**)–X–X–X–Sp–X
PSKs	
Src family/Abl/c-Fps	I–Y–G/E
Syk	(E/D)–Y–E
EGF/PDGF/IGF1/Insulin receptor PKs	E–Y–M/F/V

[a] Abbreviations: PKA, cAMP-dependent protein kinase; PKG, cGMP-dependent protein kinase; MAPKAPK1, 2, MAP kinase-activated protein kinase-1, -2; CaMK, calmodulin-dependent protein kinases; AMPK, AMP-activated protein kinase; HSV-PK, herpes simplex virus protein kinase; PRV-PK, pseudorabies virus protein kinase; ERK, extracellular signal-regulated kinase; CK2, protein kinase CK2 (casein kinase 2); G-CK, Golgi casein kinase; CK1, protein kinase CK1 (casein kinase 1); GSK3, glycogen synthase kinase-3.

[b] Amino acids are indicated by the one-letter code; B stands for any hydrophobic amino acid and X for any residue; Sp and Tp denote phosphoserine and phosphothreonine, respectively. Interchangeable residues at a given position are grouped between parentheses, and separated by slashes. The target residues are in bold type. For other sequences and/or additional information see ref. 1.

3. Design of synthetic peptide substrates

The 'ideal' peptide substrate for a given protein kinase should be:

- readily phosphorylated by the desired kinase, with favourable kinetic constants
- poorly phosphorylated, if at all, by other kinases
- suitable for a fast, simple and sensitive assay

Peptides that are efficiently phosphorylated and suitable for convenient assays can be prepared for the majority of known protein kinases, either PSKs

or PTKs. In contrast, strict selectivity is a property of only a small minority of peptide substrates. Most peptides are phosphorylated by a variety of more or less related protein kinases, albeit with different efficiencies, such a promiscuity being especially pronounced among peptide substrates of PTKs.

3.1 Efficiency of peptide substrates

In general, the design of a peptide substrate firstly requires the identification of phosphoacceptor site(s) affected by the kinase of interest in its protein targets. Often, this analysis also discloses conserved features that may represent specificity determinants. The actual relevance of these structural elements is then checked by synthesizing peptides where these residues are either conserved or substituted, and by using them as substrates for the kinase of interest. The minimum length of the peptide, required for efficient phosphorylation should also be checked. Although very short peptides (6–7 residues long) may occasionally prove excellent substrates, in general a more extended sequence is required, composed of 10–15 residues. A 'good' peptide substrate is expected to display a K_m in the low μM range (<100 μM) and a V_{max} value comparable to that of the protein substrate. One should also bear in mind that the sequence derived from the natural protein substrate is not necessarily the optimal one. Often the phosphorylation of a peptide substrate can be substantially improved by the inclusion of additional positive determinants that may be lacking in the original site, or by the elimination of negative determinants that may be present. Useful hints about substitutions or modifications likely to improve the phosphorylation efficiency can be provided by the comparative analysis of many natural phosphoacceptor sites, and by systematic studies with numerous peptide substrate derivatives.

A short-cut towards the generation of optimal peptide substrates has been recently made possible by the development of peptide library approaches. These methods, dealt with in Chapter 16, are intended to provide through a single experiment (at least in principle) an amount of information that would otherwise require hundreds of experiments with a wide variety of successively designed peptide substrates. They may also prove helpful for designing peptide substrates for kinases whose physiological phosphoacceptor sites are still unknown, and perhaps even for protein kinases which are inactive on peptides reproducing their natural substrate but might still act on unrelated short sequences. The most successful of these methods are based on oriented libraries (10, 11; see Chapter 16) consisting of peptides including either a serine or a tyrosine residue embedded in a sequence of degenerate residues. After phosphorylation by the kinase and removal of the dephosphorylated peptides, the degenerate phosphorylated peptides are sequenced by Edman degradation as a mixture. If certain amino acids are prevalent in a particular sequencing cycle, this means that the kinase had selectively phosphorylated peptides containing that amino acid at that position. In principle, one might

imagine that the sequence reconstructed from the residues most highly selected at each position would represent the optimal sequence recognized by the kinase. However, it should be remembered that this represents a *virtual* sequence. Even if the library contained a peptide with this actual sequence, the signal from a single peptide would be far too low to be detectable. The peptide library approach can have shortcomings if the importance of a residue at a given position depends on the nature of the surrounding sequence.

The potential and limitations of this approach are discussed further in Chapter 16. From the standpoint of those searching for an optimal peptide substrate, it is crucial to check if a *real* peptide designed using the *virtual* optimal sequence selected from the library is efficiently phosphorylated by the kinase. In many instances this has proved to be the case. Biases can result, however, from the tendency of oriented libraries to select specificity determinants at positions where they are not really needed, as discussed in ref. 1 (this could even worsen selectivity, see below). Other problems include the constraints imposed by the fixed scaffold of all the library peptides, notably the invariant position of the phosphoacceptor residue in the centre of the sequence, and a basic triplet, usually KKK, inserted at the C-terminal end for technical reasons (10).

3.2 Specificity of peptide substrates

Conferring absolute specificity to a peptide substrate for an individual kinase, to such an extent that no other kinase will affect it, is in most cases an unachievable goal. In many instances, however, it is possible to generate relatively selective peptides which are phosphorylated by the desired kinase much more efficiently than by the majority of other kinases, including those belonging to the same specificity class (see *Table 1*). In this connection, one can consider the example of the three types of casein kinases, all equally active on casein (hence their name) but now routinely monitored using highly specific peptide substrates (see *Table 2*). These kinases are all *acidophilic*, so-called because the specificity determinants are acidic residues surrounding the phosphoacceptor site. The design of specific peptide substrates was made possible by detailed knowledge of the specificity determinants of CK1 (requiring at least four Asp residues between n–3 and n–6), CK2 (requiring either Asp or Glu at position n+3 and additional acids nearby) and G-CK (needing Glu, but not Asp, at n+2). The n–X (n+X) nomenclature [P–X (P+X) is used by some workers] refers to the position X residues N-terminal (C-terminal) to the phosphorylated amino acid. The detailed knowledge of the specificity of the casein kinases allowed the design of peptides that fulfil the requirements of each of them, without being appreciably phosphorylated by the others, nor by basophilic or proline-directed protein kinases.

To some extent the attainment of specific peptide substrates is facilitated by incompatibilities between the specificity determinants of different kinases.

Table 2. A list of synthetic peptide substrates used for assaying protein kinases[a]

Peptide[b]	Parent protein[c]	PK[d]	K_M (μM)[e]	Ref.
LRRASLG	Pyruvate kinase	PKA	6	45, 46
		PKG	210	47
RRRASVA	Pyruvate kinase	PKA	6	48
		PKC	140	

most highly recognized by ients a *virtual* ual sequence, tectable. The e of a residue luence.

ed further in timal peptide ig the *virtual* phorylated by ues can result, cificity deter- ssed in ref. 1 s include the tides, notably centre of the minal end for

ridual kinase, cases an un- to generate esired kinase cluding those nnection, one s; all equally l using highly *idophilic*; so- rrounding the es was made unts of CK1 52 (requiring) and G-CK clature [P–X s N-terminal nowledge of les that fulfil osphorylated es.

facilitated by rent kinases.

Table 2. A list of synthetic peptide substrates used for assaying protein kinases[a]

Peptide[b]	Parent protein[c]	PK[d]	K_M (μM)[e]	Ref.
LRRASLG	Pyruvate kinase	PKA	6	45, 46
		PKG	210	47
RRRASVA	Pyruvate kinase	PKA	6	48
		PKC	140	48
		PRV-PK	20	48
		HSV-PK	n.r.	48
GRTGRRNSI	PKI	PKA	0.1	47
		PKG	2	47
GSRRRRRRY	Galine	PKC	12	49
		HCR	60	50
		dsl	300	50
QKRPSQRSKYL	MBP	PKC	7	51
KRAKRKTAKKR	MLC	PKC	0.5	52
RFAVRDRMQTVAVGVIKAVDKK	eEF-1α	PKC δ	23	53
AKRKRKGSFFYGG	(library)	PKC δ	1	54
AALVRQMSVAFFFK	(library)	PKC μ	9	54
RPRTSSF		PKB α	5	55
		$p90^{s6k}$	12	55
		$p70^{6sk}$	125	55
RPRAATF		PKB α	25	55
		$p90^{s6k}$	> 500	55
		$p70^{6sk}$	> 500	55
HMRSAMSGLHLVKRR	ACoAC	AMPK	26	56
		SNF1	108	56
AMARAASAAALARRR		AMPK	10	56
		SNF1	650	56
		CaMK I	n.r.	56
LKKLTRRPSFSAQ	ADR1	AMPK	4	57
		PKA	n.r.	57

Table 2. *Continued*

Peptide[b]	Parent protein[c]	PK[d]	K_M (μM)[e]	Ref.
LRRRLSDANF	Synapsin I	CaMK I	4	58
		CaMK II a	45	59
KKRAARATSNVFS	MLC	*CaMK IV*	0.2	59
KKRARAATSNVFA	MLC	sm MLCK	12	60
		Twitchin	4	61
PLARTLSVAGLPGKK	GS	sm MLCK	20	61
		CaMK II	54	61
KSDGGVKKRKSSSS	CaM-K II	*CaMK IV*	1	59
KKPLNRTLSVASLPGL-amide	GS	CaMK II	9	59
		CaMK IV	8	62
KKFNRTLSVA	GS	MAPAKAP-Kin 2	12	63
		CaMK II	8	63
KKLNRTLTVA	GS	$p90^{S6k}$	36	63
KKKNRTLSVA	GS	MAPAKAPK2	9	63
		CaMK II	n.r.	63
KKRNRTLSVA	GS	$p90^{S6k}$	n.r.	63
		CaMK II	n.r.	63
KKRNRTLTVA	GS	$p90^{S6k}$	0.2	9
		$p70^{S6k}$	3	9
PLRRTLSVAA	GS	$p70^{S6k}$	1	9
AVAAKKSPKKAKKPA		$p90^{S6k}$	0.7	9
NFKTPVKTIR	GS	$p70^{S6k}$	5	9
AcetylSPGRRRRKamide		$p90^{S6k}$	40	9
VTPRTPPPRR	CK2 β sub	CaMK II	4	64
		Cdc2	3	65
APRTPGGRC-polylys	MBP	Cdc2	40	66
KRELVEPKTPSGEA	EGF-R	Cdc2	1	67
		Cdk2	n.r.	68
		ERK	n.r.	68
		ERK1	74	69
		--	190	70

NFKTPVKTIR				
AcetylSPGRRRRKamide	CK2 β sub	Cdc2	40	66
VTPRTPPPRR		Cdc2	1	67
	MBP	Cdc2	n.r.	68
		Cdk2	n.r.	68

APRTPGGRC-polylys		ERK	74	69
KRELVEPKTPSGEA	EGF-R	ERK1	190	70
		ERK2	270	70
KKKKEHQVLMKTVCGTPGY	CaM-K IV	Cer.act.PK	2200	71
KKKKDPGSVLSTACGTPGY	CaM-K I	CaMKK α	263	72
KRELVEPLTLPSGEA [7]	EGF-R	CaMKK α	n.r.	72
KYHGHTMSDPGVSYR	PDH	Cer.act.PK	400	71
PLSRTLSVRSL [7]	GS	PDHK	460	73
		PhK	810	74
KRKQISVRGSL	Ph	PKA	n.r.	74
		PhK	900	64
RRREEEEESAAA		PKA	n.r.	64
		βARK	1100	75
RRRAAAAASEEE		*RhK*	2000	75
		RhK	1800	75
PRRRRRSSRPVR	Thynnine	CK2	n.r.	75
KRREILSRRPSpYR	CREB	PRV-PK	20	76
		GSK3 α	140	77
PRPASVPPSPSLSRHSSPHQSpEDEEEP [7]	GS	GSK3 β	200	77
PLSpRTLSVASLPGL-amide *	GS	GSK3	2	28
SpSpSpEESIT *	β-Casein	CK1	7	78
IGDDDDAYSITA *	Inhibitor-2	CK1	14	79
RRKHAAIGDDDDAYSITA	Inhibitor-2	CK1	58	18
RRKDLHDDEEDEAMSITA	Inhibitor-2	CK1	250	18
RRRADDSDDDDD		CK1	172	16
		CK2	19	16
RRRAEESEEEEE		G-CK	3062	17
		CK2	19	16
KKIEKFQSEEQQQ		G-CK	15	17
YIYGSFK *	β-Casein (library)	G-CK	663	17
		c-Src	55	80
		Lyn	n.r.	80
		Fyn	n.r.	80

Table 2. *Continued*

Peptide[b]	Parent protein[c]	PK[d]	K_M (μM)[e]	Ref.
AEEEIYGEFEAKKKK	(library)	c-Src	33	11
KVEKIGEGTYGVVYK	cdc2	c-Src	101	81
		Lyn	277	81
		Lck	133	81
		Fyn	487	81
EDNEYTA *	c-Src	c-Fgr	80	12
		Syk	1500	12
RRLIEDAEYAARRG		Syk	58	82
c(EDNEYTA) *	c-Src	c-Fgr	430	82
FGEFGEFGEYGEFGEFGD *	c-Src	Lyn	3800	82
KKKEEEEEEEYMPMEDL	mT	c-Src	2800	83
EAIYAAPFAKKK	(library)	Lyn	20	84
EQEDEPEGDYEEVLE *	HS1	c-Src	73	83
		v-Src	97	85
EDENLYEGLNLDDCSMYEDI *	ARAM Igα	v-Abl	97	85
		c-Abl	4	11
EEEPQYEEIPIYLELLP *	mT	Syk	4	12
		Csk	625	12
EEEPQFEEIPIYLELLP *	mT	Lyn	5000	12
		Syk	25	12
KKSRGDYMTMQIG	IRS-1	c-Fgr	33	12
		Lyn	98	12
		Csk	211	12
		Csk	34	2
		c-Fgr	185	2
		Csk	63	2
		c-Fgr	473	2
		IR	30	86
		IGF1 R	178	86

		Csk	34	2
EEEPQFEEIPIYLELLP *	mT	c-Fgr	185	2
		Csk	63	2
KKSRGDYMTMQIG	IRS-1	c-Fgr	473	2
		IR	30	86
		IGF1 R	178	86

		v-Src	300	85
		v-Abl	520	85
KKKSPGEYVNIEFG	IRS-1	IGF1 R	26	86
		IR	52	86
DRVYIHPF *	Angiotensin II	Lyn	1200	82
		several PTKs		

[a]A selection of peptide substrates with K_m values in the lowest range for the indicated kinases are reported. In some cases, the peptides have been chosen for their more favourable V_{max}/K_m ratio, even if their K_m values are not the lowest reported. For a detailed analysis and comparison with other peptides, see the indicated references.

[b]Sequences are indicated with the one-letter code for amino acids; the target residues are bold type; Sp, phosphoserine; an asterisk (*) or a question mark (?) denote peptides whose phosphorylation assay by the phosphocellulose paper method is inapplicable or dubious, respectively.

[c]The parent protein is indicated, when applicable, even if the peptide sequence does not correspond exactly to the original one because of substitutions or additions introduced to improve phosphorylation and/or specificity, or to allow phosphocellulose paper assay.

Abbreviations: ACoAC, acetyl–CoA carboxylase; ADR1, yeast transcriptional activator of the *ADH2* gene; ARAM, antigen recognition activation motif; CREB, cAMP response element binding protein; GS, glycogen synthase; HS1, haematopoietic lineage-cell specific protein; IRS-1, insulin receptor substrate-1; MBP, myelin basic protein; MLC, myosin light chain; mT, middle T; PDH, pyruvate dehydrogenase; Ph, phosphorylase; PKI, heat stable inhibitor protein of PKA.

[d]Some protein kinases (denoted in *italics*) are highly specific enzymes whose peptide substrates are phosphorylated very slowly despite their sometimes favourable K_m values.

Abbreviations: AMPK, AMP-activated protein kinase; βARK, β adrenergic receptor kinase; CaMK, calmodulin-dependent protein kinases; Cdc2 kinase, protein kinase expressed by *CDC2* gene; cdk, cyclin-dependent kinase; Cer. Act. PK, ceramide activated protein kinase; CK1, protein kinase CK1 or casein kinase 1; CK2, protein kinase CK2 or casein kinase 2; Csk, C-terminal Src protein kinase; dsI, double strain RNA inhibitor = eIF2α kinase; ERK, extracellular signal-regulated kinase; G-CK, Golgi casein kinase; GSK3, glycogen synthase kinase-3; HCR, heme controlled repressor = eIF2α kinase; IGF1 R, insulin-like growth factor-1 receptor; I-R, insulin receptor; MAPKAP, MAP kinase activated protein kinase; PDHK, pyruvate dehydrogenase kinase; PhK, phosphorylase kinase; PKA, cAMP-dependent protein kinase; PKB = RAC, PK related to PKA and PKC; PKG, cGMP-dependent protein kinase; PRV-PK, pseudorabies virus protein kinase; RhK, rhodopsin kinase; sm MLCK, smooth muscle myosin light chain kinase; SNF1, sucrose non-fermenting protein kinase.

Thus, a proline at n+1, which is the *sine qua non* for of all proline-directed kinases, acts as a powerful negative determinant for the majority of other kinases (1). Similarly, the basic residues that are specificity determinants for *basophilic* kinases, and also several proline-directed kinases, are detrimental to acidophilic kinases like CK2.

Another tool to improve the selectivity of peptide substrates is replacement of the phosphoacceptor residue: this may prove helpful when two kinases with otherwise similar consensus sequences display a markedly different tolerance for threonine versus serine (9).

In the case of PTKs, due to their promiscuity and lack of stringent site specificity, the generation of selective peptide substrates is particularly difficult. It appears that the positions where specificity determinants act more potently in the case of PTKs are those flanking the tyrosine at n–1 and n+1 (see *Table 1*). A bulky hydrophobic residue at n–1 (generally Ile) is strongly selected in an oriented library by all the Src family kinases tested so far, as well as by Abl, Csk ,and c-Fps (2, 11), a predilection that has been confirmed with conventional peptide studies (2). In contrast, receptor kinases (and Syk) display a marked preference for acidic residues (generally Glu) at n–1 (11, 12). The nature of the preferred residue at n+1 appears to be dictated by the composition of the so-called 'P+1 loop'. In cAMP-dependent protein kinase (PKA) this loop forms a hydrophobic pocket, with the last residue of the pocket, Leu-205, making the main contact with the determinant at position n+1. In PTKs the residue equivalent to Leu-205 in PKA is either methionine, indicating a preference for a hydrophobic residue at n+1 (e.g. receptor PTKs and Csk), or threonine, which is compatible with the accommodation of glycine or another hydrophilic residue (e.g. the Src family and other non-receptor kinases). An exception is provided by Syk, where the equivalent residue to Leu-205 in PKA is a tyrosine. Syk displays a striking preference for peptides where tyrosine is embedded between several acidic residues. Based on this principle, excellent peptide substrates for Syk could be designed (12) (see *Table 2*). It is not known how efficiently these peptides are phosphorylated by the Syk-related ZAP kinase. Although some of these highly acidic peptides display poor kinetics with other non-receptor PTKs (e.g. 12), they are still appreciably phosphorylated by the latter, so these peptides are not entirely specific for Syk. If, however, they are used in combination with different peptide substrates (e.g. the Cdc2-derived peptide, see *Table 2*) the activity ratio can be used as a reliable criterion to identify Syk activity.

A list of peptide substrates that have been used for the assay of a variety of protein kinases is provided in *Table 2*, with the indication of K_m values, whenever available. A number of points should be made. First, the number of protein kinases matching a given peptide does not allow any direct inference about actual specificity, since the list is incomplete and does not include in any case those kinases that have been shown to be *inactive* on the peptide considered. For a deeper insight into peptide selectivity, it is advisable to consult

the pertinent reference(s). Secondly, in many instances differences in K_m values can be exploited to improve the selectivity of peptide substrates towards certain kinases, by performing the assay with appropriate low peptide concentrations. Thirdly, a low K_m does not always indicate a selective substrate, and *vice versa*. The outstanding affinity of the peptide GRTGRRNSI for PKA (K_m = 0.1 μM) does not make it more specific than the traditional 'kemptide', LRRASLG (K_m = 6 μM); both peptides are phosphorylated by several basophilic protein kinases. By contrast, the β-casein-derived peptide KKIEKFQSEEQQ is a very specific substrate for the Golgi apparatus casein kinase (G-CK), despite its high K_m (663 μM), since it is not appreciably acted upon by other protein kinases. Sometimes, an increase in K_m value is acceptable if it is the consequence of a modification (such as an addition of basic residues) aimed at rendering the peptide suitable for a fast and reliable assay technique (see Section 3). In this connection it should be noted that the majority of the peptides listed in *Table 2* either are (or have been made) suitable for the phosphocellulose paper assay; exceptions are indicated by an asterisk or question mark.

Clearly, a better knowledge of negative determinants variably perceived by different kinases will prove helpful for tailoring peptides that display higher selectivity. Another strategy that might deserve attention is to exploit the *active-site* specificity with artificial phosphoacceptor derivatives different from the natural hydroxyl amino acids, i.e. serine, threonine, and tyrosine. Especially promising in this respect appear to be studies with alcohol-bearing compounds that can serve as protein kinase substrates, and which are disclosing unexpected differences in the active-site specificity of otherwise closely related kinases, both PSKs and PTKs (13, 14). Likewise, the use of tyrosine analogues could provide a key for monitoring individual members of the PTK family. Pertinent to this is the observation that the replacement of tyrosine by its analogue, 7-hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, in a cyclic peptide derived from the activation loop of c-Src fully prevents phosphorylation by the Src-related kinase c-Fgr, while it is tolerated fairly well by Syk (Ruzza, P., Donella-Deana, A., Calderan, A., Filippi, B., Cesaro, L., Pinna, L.A., and Borin, G., unpublished data).

3.3 Measurement of peptide phosphorylation

Even the best and most specific peptide substrate would be of little use unless a reliable, sensitive, and simple method for assaying its phosphorylation was available. If phosphorylation is performed by radiolabelling with [γ-^{32}P]ATP, it is normally sufficient to achieve a fast and complete separation of the radiolabelled phosphopeptide from the large excess of radioactive ATP, and its hydrolysis product [^{32}P]phosphate. In an assay where both radioactive ATP and the phosphoacceptor peptide substrate are present at 500 μM, and assuming that only 5% of the peptide is converted into its phosphorylated form, at the end of incubation there will be an ~20-fold excess of radioactive

ATP over radioactive phosphopeptide. In order for the level of residual radioactivity due to ATP to be acceptable (e.g. <10% of the radioactivity incorporated into the phosphopeptide), no more than 0.5% of the original ATP would remain in the sample after the separation procedure. This is an optimistic scenario when the kinase is poorly active, or when the kinetics are performed with very low concentrations of peptide. In these cases it is desirable that <0.01% of the unreacted ATP remains to contaminate the phosphopeptide product, in order to make the assay feasible and reliable.

This performance can be achieved using the phosphocellulose paper procedure, based on the principle that at low pH values ATP (and inorganic phosphate) does not bind to phosphocellulose, whereas any singly phosphorylated peptide including two or three basic residues will bind. This method, first developed by Glass *et al.* (15), can be directly applied to most peptide substrates of basophilic protein kinases, which already include basic residues as specificity determinants (see *Table 2*). It can also be adapted to peptide substrates for kinases that, while not requiring basic residues, nevertheless tolerate them. This includes proline-directed kinases, and many PTKs. Although it is commonly stated that the latter require acidic determinants, in reality they are quite promiscuous as far as peptide substrates are concerned and, with a few exceptions, will phosphorylate peptides to which basic residues have been added quite well.

To ensure that phosphopeptides quantitatively bind to phosphocellulose paper, it is advisable to include at least three basic residues rather than two (see also Chapter 9 for a discussion of this topic). An example of this is provided by angiotensin II, an octapeptide widely used for the assay of PTKs, which has only two basic residues (DRVYIHPF). In our experience, angiotensin II itself is not suitable for reliable assays, since 30–40% of its phosphorylated derivative can be lost upon washing of phosphocellulose paper (see *Table 3*). To make binding quantitative (> 95% recovery after washing) another arginine can be added at the C-terminal end. This additional arginine did not appear to impair phosphorylation, at least by Src-related kinases.

Even three basic residues may not be sufficient if the peptide is phosphorylated at more than one site. This may be the case if more than one phosphoacceptor residue is present, or if additional phosphorylated residue(s) are included from the outset to prime phosphorylation by 'phosphate-directed' kinases (e.g. GSK3, see *Table 3*).

Whenever the basic residues are not present in the 'natural' sequence which formed the basis for the peptide substrate, they must be inserted at position(s) where they do not compromise phosphorylation efficiency or selectivity [e.g. by rendering the peptide a substrate for undesired basophilic kinase(s)]. This is exemplified by the design of peptide substrates suitable for the phosphocellulose paper assay of casein kinases (16, 17). These kinases are acidophilic and often perceive basic residues as negative determinants. In the case of CK2, whose specificity determinants are mostly located C-terminal to the

Table 3. Variable recovery of phosphopeptides by different procedures[a]

Peptides	Recovery of phosphopeptide by:[b]		
	P-cellulose paper	AG 1-X8 anion exchange	DEAE-paper
DRVYIHPF (angiotensin II)	45	100	n.t.
DRVYIHPFR	95	100	n.t.
EDNEYTA	< 1	90	n.t.
RRLIEDNEYTARG	98	98	n.t.
SAEEEDQYN	< 1	82	n.t.
RRRADDSDDDD	97	30	80
ESEEEEE	< 1	30	82
RKMKDTDSEEEIR	98	85	n.t.
RRASVA	98	98	n.t.
SpSpSpEESIT	< 1	< 1	n.t.
EQEDEPEGDYEEVLE	n.t.	60	88
DEDADIYDEEDYDL	n.t.	20	97
PEGDYEEVLE	n.t.	70	70
PEGDYAAVLE	n.t.	98	20

[a]Unpublished data provided by A. Donella Deana and F. Meggio.
[b]Values are expressed as a percentage relative to the recovery of phosphopeptide by the acid hydrolysis method, assumed to be 100%; n.t., not tested.

serine, the problem could be circumvented by placing the basic residues upstream at positions n–4 to n–6, where they are reasonably well tolerated. At position n–1 and anywhere downstream they are not tolerated. The resulting peptide, RRRADDSDDDDD (*Table 2*), besides being an excellent substrate for CK2 by virtue of its seven Asp residues, and suitable for the phosphocellulose paper method by virtue of its three N-terminal arginines, is also quite specific. The lack of basic residues at n–2 and n–3 prevents its phosphorylation by most basophilic kinases, whereas the alanine at n–3 (instead of another Asp) prevents phosphorylation by another casein kinase, CK1. The Golgi apparatus casein kinase (G-CK) , whose consensus sequence is S-X-(E/Sp)-X, does not phosphorylate this peptide because it contains Asp rather than Glu at n+2 (see *Table 1*). Conversely, a highly specific peptide substrate for G-CK, suitable for the phosphocellulose assay, was derived from the sequence surrounding serine-35 in β-casein (KKIEKFQSEEQQQ) exploiting the basic residues present in the N-terminal moiety of the natural sequence.

In the case of CK1, which requires either a phosphorylated residue at n–3, or an acidic cluster upstream from n–2, placing the basic residues downstream was not successful since the resulting peptide (DDEEDEEMSETARRR) was a very poor substrate (18). This problem was circumvented by placing the basic triplet upstream, quite far from the acidic cluster. The resulting peptide (RRKDLHDDEEDEAMSITA), although having a K_m higher than that of the peptide devoid of basic residues, is still a fairly good substrate (16).

If the phosphocellulose paper method is not applicable, alternative

procedures (described below) based on AG 1-X8 anion-exchange assay and DEAE-cellulose paper are available. The former is not successful with acidic peptides, which tend to be retarded and elute more or less together with the radioactive ATP. It would be advisable in these cases to replace acidic residues with neutral ones if they are not required as specificity determinants. The opposite strategy applies to the DEAE-cellulose paper method, which exploits the fact that at neutral pH (where aspartic and glutamic acid side chains are completely dissociated) very acidic peptides bind much more tightly than ATP, which can be eluted prior to the peptide. This method is quantitative only if the peptide contains at least seven net negative charges (see *Table 3*), so its applicability can be improved by attaching additional acidic residues, provided that they are not detrimental to phosphorylation efficiency. In this respect, DEAE-cellulose paper is complementary to phosphocellulose paper, and can be applied to peptide substrates for kinases that do not tolerate basic residues, while accepting a large number of acidic residues. It has been successfully applied to CK2 assays (19) and to the PTK Syk (see *Table 3*), both of which are acidophilic. This method may also prove useful with other acidophilic kinases, such as CK1 and GSK3.

There is also a procedure that works with any kind of phosphopeptide, irrespective of the amino acid composition and the identity of the phosphoamino acid. This is based on acid hydrolysis of ATP under conditions where the phosphoester bonds of the peptides are not affected. Radioactive inorganic phosphate is removed by molybdate treatment and extraction of the phosphomolybdate complex with an organic solvent (see below). A limitation of this 'universal' method, as well as the fact that it is more complicated than the paper methods, is the high background. This is due to incomplete removal of [^{32}P]phosphate generated from [γ-^{32}P]ATP hydrolysis, which is a particular problem if the kinase activity to be measured is low. If this is not the case, the method may be the first choice procedure to compare the phosphorylation of peptides of markedly different composition (e.g. basic, acidic, and neutral), which could not be compared using a single ion-exchange procedure. The method can also be used to check the applicability of the other methods (as in *Table 3*).

4. Phosphorylation of peptides

4.1 The phosphorylation reaction

4.1.1 Assay conditions

Assays are usually performed at 30°C in a total volume of 20–50 µl. Typical final concentrations in the assay are:

- 20–50 mM buffer (usually Tris–HCl or Hepes) at pH 7.0–8.0
- 10–500 µM [γ-^{32}P]ATP (specific radioactivity 500–5000 cpm/pmol)
- 5–10 mM $MgCl_2$ or $MnCl_2$, or both

Other components to be added are the peptide substrate, the kinase, and any essential co-factor such as an allosteric activator of the kinase. The optimal conditions depend on the specific enzyme being used.

ATP is utilized by all protein kinases as a phosphate donor, but some (e.g. CK2, and to a lesser extent Cdc2) can use GTP as well. This can form the basis of a more specific assay for these kinases. An ATP concentration of 10 μM is usually sufficient for straightforward activity assays, but higher concentrations are required for quantitative analysis, especially the determination of K_m values for peptides. In this case, the ATP concentration should exceed the K_m value for ATP (for most kinases this is in the range 1–150 μM) by several-fold. The desired specific radioactivity of [γ-^{32}P]ATP can be obtained by mixing non-radioactive ATP with the required amounts of [γ-^{32}P]ATP from a 'carrier free' stock (for example, the solution in 50% ethanol supplied by Amersham at a specific radioactivity of 3000 Ci/mmol). ^{32}P has a half-life of only 14.3 days; decay tables are available to calculate the specific activity on the day of use. A precise assessment of [γ-^{32}P]ATP specific activity is essential for calculating reaction velocities and the stoichiometry of phosphorylation (see *Protocol 2* in Chapter 9).

Mg^{2+} or Mn^{2+} ions are required both to form the $Mg–ATP^{2-}$/$Mn–ATP^{2-}$ complex which is the true substrate of all protein kinases, and because binding of divalent metals may also be required at other sites on the enzyme. Usually Mg^{2+} is preferred by PSKs, and Mn^{2+} by PTKs. However, many exceptions to this are known, and a mixture of both ions is often used. In principle, the use of either Mg^{2+} or Mn^{2+} could improve the selectivity of the assay.

Protocol 1. Standard peptide phosphorylation reaction

Equipment and reagents

- thermostatic water bath or incubator
- vortex mixer
- microcentrifuge
- equipment for radioactive safety (e.g. Perspex shields, Geiger counter)
- assay buffer (0.5 M Tris–HCl or Hepes–NaOH at pH 7.0–8.0)
- [γ-^{32}P]ATP (3000 Ci/mmol; Amersham supplies a stock in 50% ethanol)
- $MgCl_2$ and/or $MnCl_2$ (100 mM)
- unlabelled ATP (50 mM; make up in water and neutralize with Tris base or dilute NaOH)
- peptide solution (10 mg/ml; most peptides are soluble in water, but neutralize with Tris base or ammonia if necessary)
- kinase
- appropriate co-factors/activators

Method

1. A mix of the reagents which are common to all of the assays can be prepared in advance. Add the reagents (buffer, [γ-^{32}P]ATP, cold ATP, $MnCl_2$, and/or $MgCl_2$, peptide substrate, co-factors) to polypropylene tubes. Adjust to the desired final volume with water. Vortex and centrifuge for few seconds.
2. Dilute the kinase in the appropriate buffer. Start the reactions by

Protocol 1. *Continued*

adding the enzyme to each tube at fixed intervals of time; mix and place the tubes in the incubator at 30 °C.

3. Stop the reactions by one of the methods described in Section 4.2.
4. Blank reactions should contain all reagents except the kinase; if a significant autophosphorylation of the kinase occurs, an additional blank should contain all components except the peptide substrate.

The peptide substrate can be initially used at a wide range of concentrations (0.1–2 mM); a kinetic analysis can then be performed to determine the most appropriate concentration to be used in routine assays. In some instances, 2-mercaptoethanol (10 mM) or dithiothreitol (0.1 mM) is added to maintain proteins in the reduced state, while BSA (1 mg/ml) and/or a detergent [such as Triton X-100, or Tween 80 (0.02–0.1% v/v)] are added to stabilize the kinase. The detergent may also help to avoid binding of the peptide to the assay tubes. The appropriate ionic strength should be determined by initial trials, since many protein kinases are inhibited by high salt, while others require a defined salt concentration for optimal activity. Reactions are typically carried out for 5–15 minutes, after which they are stopped in a manner depending on the method of analysis (to be discussed in later sections). *Protocol 1* describes a typical kinase assay reaction. See *Protocol 1* in Chapter 9 for a variation on the same theme.

4.2 Techniques for post-assay separation of phosphopeptide and ATP

To detect peptide phosphorylation, it is necessary to separate the radioactive phosphopeptide from unhydrolysed [γ-^{32}P]ATP, and [^{32}P]phosphate which may also be present due to the presence of ATPases or phosphatases (Section 3.3). The technique employed to this aim depends on the nature of the peptide being tested. Several different methods are available, as described in the following sections.

4.2.1 Phosphocellulose paper assay

This technique (15) is the most convenient for basic peptides; it exploits the binding of the peptide to phosphocellulose cation-exchanger paper, under conditions of low pH which protonate the negative charges due to carboxylic side chains and prevent binding of [γ-^{32}P]ATP (see *Protocol 2*). A version of this technique is also given as *Protocol 1* in Chapter 9.

The criteria necessary for a peptide to bind quantitatively to phosphocellulose paper under these conditions are discussed in Section 3.3. Whenever assays are conducted for the first time with a new peptide, trials should be performed using different techniques of separation of peptide and ATP, as shown in *Table 3*.

Protocol 2. Kinase assay using phosphocellulose cation-exchange paper

Materials

- phosphocellulose cation-exchange paper (P81 Whatman; cut into 2x2 cm squares)
- phosphoric acid (0.5 %, v/v)
- hair dryer
- liquid scintillation counter
- 500 ml plastic beaker (for washing papers; this beaker should have inserted into it a slightly smaller plastic beaker, in which holes of about 0.5 cm diameter have been made, to retain the papers during washing)

Method

1. After the desired incubation time (see *Protocol 1*), stop the phosphorylation reaction by spotting the reaction mixture (or an aliquot of it) on to the P81 paper squares.
2. When the liquid is adsorbed, drop the papers into the 500 ml beaker, which contains about 200 ml of phosphoric acid. Place the beaker on to a magnetic stirrer for about 5 min; discard the liquid as radioactive waste and replace it with 200 ml of fresh phosphoric acid.
3. Wash three times with phosphoric acid (200 ml per wash for ~5 min each time).
4. Wash for a few seconds with acetone and dry papers with a hair dryer.[a]
5. Place the paper squares into scintillation vials and count them in scintillation fluid.

[a] This step is not essential, but the paper squares are easier to handle when dry.

4.2.2 AG 1-X8 anion-exchange chromatography assay

For phosphopeptides which do not contain enough basic residues to allow binding to phosphocellulose paper, separation of [γ-^{32}P]ATP can be achieved by chromatography on AG 1-X8 anion-exchange resin (20). In 30 % (v/v) acetic acid, ATP binds quite strongly to the resin by virtue of its three negative charges, while phosphopeptides bearing just one phosphate group usually do not. They pass through the column and can be used for further analysis. They can be directly counted for incorporated radioactivity by Cerenkov counting (without addition of scintillant), and can be recovered by lyophilization for further use (e.g. as a radioactive substrate for protein phosphatase, see Section 5.1.1 and *Protocol 7*).

More often, a combination of AG 1-X8 chromatography and acidic hydrolysis followed by phosphate extraction (see Section 4.2.3) is a good device to ensure that the measured radioactivity is incorporated into the phosphopeptide and not due to contaminating [^{32}P]phosphate or [γ-^{32}P]ATP (see *Protocol 4*). If the peptide contains acidic residues, it will bind to the column

more tightly and may be difficult to resolve completely from [γ-^{32}P]ATP. For this reason, the AG 1-X8 anion-exchange chromatography assay is not suitable for very acidic peptides, especially multiply phosphorylated peptides (see *Table 3*).

4.2.3 Assay by acidic hydrolysis of [γ-^{32}P]ATP and separation of ^{32}Pi from phosphopeptide

If the phosphorylated peptide contains a similar number of net negative charges to ATP in 30% acetic acid, it is not suitable for analysis by AG 1-X8 chromatography. In these cases, the first choice method is the hydrolysis of [γ-^{32}P]ATP, followed by extraction of the phosphate (as a phosphomolybdate complex) in an organic phase (21). Treatment at 100°C for 15 minutes in the presence of 1 N HCl is sufficient to hydrolyse the acid anhydride bonds of ATP, while preserving the ester bonds of phosphoamino acids. The reaction is neutralized by addition of NaOH, and the formation of a phosphomolybdate complex is induced by the presence of ammonium molybdate. The phosphomolybdate complex is soluble in organic solvents such as isobutanol/toluene (1:1) (22), allowing its extraction from the phosphopeptide, which remains in the aqueous phase. The radiolabelled phosphopeptide can be quantified by counting an aliquot of the aqueous phase by scintillation counting.

This assay method is appropriate for SerP, ThrP, and TyrP-containing peptides, since very little hydrolysis of any of these phosphoamino acids occurs under the conditions used. The method works with all peptides, no matter how acidic or basic, with the possible exception of markedly amphipathic peptides which may concentrate at the water/organic phase interface. This is, therefore, the first choice method whenever a comparison of phosphorylation of a variety of peptides of markedly different composition is being undertaken. It can be applied directly after the phosphorylation reaction (see *Protocol 3*) or, for suitable peptides, after chromatography on AG 1-X8 resin (see *Protocol 4*) in order to reduce the background, which is the main drawback of this method (see Section 3.3).

Protocol 3. Kinase assay using acid hydrolysis and phosphomolybdate extraction

Material

- HCl (2 M and 1 M)
- ammonium molybdate [5% (w/v); to 60 ml H_2O, add 22.4 ml concentrated H_2SO_4 (18 M) and 10 g ammonium molybdate (add in small amounts, waiting for solubilization before adding more); adjust volume to 200 ml with water]
- solution A[a] (mix 50 ml of 1 M NaOH, 50 ml of 5% ammonium molybdate, 38 ml of 50% TCA, and water to a final volume of 190 ml)
- NaOH (1 M)
- TCA (trichloroacetic acid, 50% w/v)
- water-saturated isobutanol/toluene (1:1) (mix 1 part isobutanol to 1 part toluene and about 0.5 part H_2O; wait for complete separation of aqueous and organic phases and take the organic phase)
- incubator at 100°C, or boiling water-bath
- vortex mixer or rotational shaker
- liquid scintillation counter

Method

1. Stop the phosphorylation reaction[b] by adding an equal volume of 2 M HCl, in order to obtain a final concentration of 1 M HCl.
2. Add 1 M HCl to a final volume of 0.5 ml.
3. Boil for 15 min.
4. Stop the hydrolysis by placing the tubes into ice and adding 1.9 ml of solution A.
5. Allow to cool, then add 5 ml of isobutanol/toluene (1:1). Cap the tubes and shake for about 30 sec.[c]
6. Wait for physical separation of the two phases; remove the upper hydrophobic layer and discard as radioactive waste.
7. Repeat steps 5 and 6 twice more.
8. Withdraw 2 ml of the lower aqueous phase and count in at least 8 ml of scintillation liquid, after vigorous shaking.
9. Calculate the total radioactivity incorporated into the phosphopeptide, multiplying the cpm detected in 2 ml of the aqueous phase by the factor 1.2.

[a] Stable for at least two weeks at room temperature.
[b] Incubations should be carried out in 10 ml tubes with caps resistant to boiling.
[c] Mixing of the aqueous and organic phases can be obtained by vortexing, or using a rotational shaker.

Protocol 4. Kinase assay by combining AG 1-X8 chromatography and acid hydrolysis

Materials

- disposable columns (e.g. Pasteur pipettes plugged with glass wool), and a suitable rack to hold them
- AG 1-X8 anion-exchange resin [equilibrated in 30% (v/v) acetic acid]
- acetic acid [30% (v/v)]
- HCl (6 M)
- ammonium molybdate [5% (w/v), prepared as in *Protocol 3*]
- solution B[a] (mix 15 ml of 12 M NaOH, 50 ml of 5% ammonium molybdate and 30 ml of 50% TCA)
- water-saturated isobutanol/toluene (1:1) (prepared as in *Protocol 3*)
- incubator at 100°C, or boiling water-bath
- vortex mixer or rotational shaker
- liquid scintillation counter

Method

1. Stop the phosphorylation reactions by adding 0.5 ml of 30% (v/v) acetic acid.[b]
2. Prepare a number of columns corresponding to the number of samples, by pouring about 0.5 ml AG 1-X8 resin into each column. Wash the resin extensively with 30% acetic acid.

Protocol 4. *Continued*

3. Load the samples on to the columns and collect the flow through fraction.
4. Wash three times with 1 ml 30% (v/v) acetic acid, and collect the eluate in the same tubes as the flow through, to a total volume of 3.5 ml.[c]
5. Withdraw 1 ml from the eluate, transfer it into 10 ml polypropylene tubes, and add 0.6 ml 6 N HCl.
6. Boil for 15 min.
7. Stop the hydrolysis by placing the tubes into ice and adding 1.9 ml of solution B.
8. Follow *Protocol 3* from step 5 to step 8.[d]
9. Calculate the total radioactivity incorporated into the phosphopeptide, multiplying the cpm detected in 2 ml of aqueous phase by the factor 6.125.

[a] Stable for at least 2 weeks at room temperature.
[b] If the incubation volume is higher than 50 µl, make sure the final concentration of acetic acid is 30% (v/v) by adding an appropriate amount of a more concentrated solution.
[c] The assay can be terminated at this step by directly counting the eluate, as described in Section 4.2.2. If high background values are observed (in the absence of enzyme or peptide), possibly due to contamination with [γ-^{32}P]ATPor the presence of other acid-labile compounds, continue the assay from step 5.
[d] Since only small amounts of acid-labile compounds are usually present after AG 1-X8 chromatography (compared with conditions described in *Protocol 3* where all of the [γ-^{32}P]ATP is hydrolysed), step 7 of *Protocol 3* can often be skipped.

4.2.4 Assay by detection of radiolabelled phosphoamino acids

Detection and quantitation of a peptide phosphorylation reaction can also be performed by the hydrolysis of peptide bonds (6 M HCl, 4 h at 110°C) followed by separation of phosphoamino acids. The technique can be successfully employed for phosphoserine- (SerP) and phosphothreonine- (ThrP) containing peptides, but it is not useful for phosphotyrosine- (TyrP) containing peptides, since the hydrolysis of this phosphoamino acid is too high under the conditions required for peptide hydrolysis. Compared with the other methods for phosphorylation assay, this one has the advantage that it allows separation between SerP and ThrP, thus giving useful information in the case of peptides containing both kinds of amino acids. The separation can be achieved by high voltage paper electrophoresis (23) (*Protocol 5*). At pH 1.9, SerP and ThrP are equally negatively charged and both migrate towards the anode, but migration of SerP is faster than that of ThrP due to lower steric hindrance. [^{32}P]Phosphate, derived from ATP hydrolysis, separates from both phosphoamino acids due to its much higher mobility. Quantitation can be achieved by means of autoradiography of the paper and scintillation counting of the spots corresponding to the radiolabelled phosphoamino acids In

quantitative (or comparative) studies, corrections must be made for hydrolytic losses of SerP (48%) and ThrP (14%) (24). The general applicability of this procedure was established by assaying the ^{32}P incorporated by different protein kinases into a series of peptides with dissimilar structures, and demonstrating that they underwent comparable loss of inorganic phosphate (25).

Alternatively, two-dimensional separation of phosphoamino acids on thin-layer cellulose plate can be performed (26) (see *Protocol 11* in Chapter 5). This is convenient when a qualitative analysis of TyrP-containing peptides is required, since TyrP is not resolved from ThrP under the conditions described in *Protocol 5*. In order to avoid complete release of free phosphate from phosphotyrosine, only a partial hydrolysis of the peptide is performed in this case (6 M HCl at 110°C for 1 h instead of 4 h).

Protocol 5. Kinase assay by peptide hydrolysis and separation of phosphoamino acids

Materials

- HCl (6 M)
- sealable glass vials
- 110°C oven
- drying device (centrifugal vacuum concentrator, e.g. Savant SpeedVac equipped with an NaOH trap to collect acid, or an air-flow device)
- pH 1.9 buffer [mix 50 ml formic acid (98%), 156 ml glacial acetic acid, and water to 2 litres]
- phosphoamino acid standards (20 mM SerP, 20 mM ThrP)
- ninhydrin solution [0.2% (w/v); dissolve 100 mg ninhydrin in 49 ml acetone and 1 ml glacial acetic acid]
- Whatman paper (3MM Chr)
- hair dryer
- high voltage electrophoresis apparatus
- equipment for autoradiography and scintillation counting

Method

1. Stop the phosphorylation reaction by transferring the mixture into sealable glass vials containing about 3 ml of 6 M HCl.
2. Seal the vials and place them in the 110°C oven for 4 h.
3. Cool the vials on ice, open them, and completely evaporate the liquid (with centrifugal vacuum concentrator or air flow).
4. Wash twice, adding about 1 ml of water and taking to dryness each time.
5. Add 50 µl of buffer at pH 1.9 and vortex thoroughly to allow complete solubilization of the dried hydrolysis products.[a]
6. Add 5 µl of cold phosphoamino acid standard solution to each sample.
7. Prepare a strip of Whatman 3MM paper with the desired width;[b] mark the origin with a pencil line a few cm from the cathode end. Mark the positions to load each sample 1–2 cm apart on the origin.

Protocol 5. *Continued*

8. Load the samples a small amount at a time, allowing to drying between each application.[c]
9. Wet the paper with pH 1.9 buffer. Avoid directly wetting the origin line, but allow the buffer to soak into it from both sides by capillary action. Absorb excess buffer from the paper surface using a blotter.
10. Place the paper strip on the high voltage electrophoresis apparatus, whose buffer tanks have been filled with pH 1.9 buffer. If necessary, connect the strip to the tanks by means of paper bridges, wetted with pH 1.9 buffer. Apply a voltage (150 V/cm of paper width[d]) for 150 min, using a suitable cooling system.
11. Stop the electrophoresis, dry the strip, and spray it with 0.2% (w/v) ninhydrin solution. Place the paper into the oven for few min. Purple spots will appear where the phosphoamino acid standards are present.
12. Detect the radiolabelled phosphoamino acids by autoradiography. Cut out the radioactive spots and count in a liquid scintillant.[e] Apply a correction for hydrolytic losses of phosphoamino acids (see Section 4.2.4).

[a] It is often necessary to repeat this step after the first loading of the sample on the paper (step 8), in order to totally remove radioactivity from the vial.
[b] Note that using a paper strip which is too wide can induce overheating. At the usual 3000 V used for electrophoresis, the maximal width is about 15 cm, but it can vary according to the cooling device used.
[c] This step can be speeded up by using a hair dryer.
[d] 3000 V is usually suitable. If resolution of SerP and ThrP is not necessary, the electrophoretic time can be reduced to 1 h, since this will allow adequate resolution from [^{32}P]phosphate.
[e] Alternatively, use a phosphorimaging system (e.g. InstantImager, Packard).

4.2.5 DEAE-cellulose paper assay

This technique (19) (*Protocol 6*) is complementary to the phosphocellulose paper assay, being suitable for assays with acidic rather than basic peptides. It exploits the binding of the peptide to DEAE- (diethylaminoethyl-) cellulose anion-exchange paper, under conditions of pH and ionic strength which do not allow binding of [γ-^{32}P]ATP. Phosphopeptides with at least six acidic amino acids are detectable by this method, but the binding is quantitative only for peptides containing eight or more acidic residues. In this latter case the results are comparable with those obtained by total acid hydrolysis (Section 4.2.3 and *Table 3*), but provide a much faster method of analysis.

Protocol 6. Protein kinase assay using DEAE-cellulose anion-exchange paper

Materials

- 0.5 M EDTA
- DEAE-cellulose paper (NA 45, Schleicher & Schuell, NH, cut into 2 cm × 2 cm squares)
- washing buffer (100 mM NaCl, 50 mM Tris–HCl, pH 7.4)
- 500 ml beaker, with a smaller plastic beaker insert (see *Protocol 2*)
- hair dryer
- liquid scintillation counter

Methods

1. Stop the phosphorylation reaction by addition of 0.5 M EDTA to obtain a final concentration of 0.14 M.
2. Spot the reaction mixture (or an aliquot of it) onto the DEAE-cellulose paper squares.
3. When the liquid is adsorbed, drop the papers into the 500 ml beaker containing about 200 ml of washing buffer. Place the beaker on to a magnetic stirrer for about 5 min; discard the liquid as radioactive waste and replace it with 200 ml of fresh washing buffer.
4. Repeat washing three times with washing buffer (200 ml per wash for at least 5 min).
5. Wash for a few seconds with acetone and dry the papers with a hair dryer.[a]
6. Place the paper squares into scintillation vials and count them in scintillant.

[a] This step is not essential, but the paper squares are easier to handle when dry.

4.2.6 Alternative methods for protein kinase assay

A number of other techniques have been employed, with different advantages and disadvantages according to the type of substrate and the phosphorylation conditions used:

- SDS–PAGE (27);
- isoelectric focusing (28);
- PEI–cellulose column chromatography (29);
- binding to ferric adsorbent paper of tritiated phosphopeptide (30) (the method is radioactive, employing ^{3}H, but avoids the use of the higher energy ^{32}P radioisotope);
- assay on streptavidin-linked disks, which can capture previously biotinylated peptides (31) (this method has the advantage of a low background in

crude extracts, since the binding of other phosphorylated, but not biotinylated, proteins does not occur);

- a thin-layer chromatography technique suitable for both acidic and basic peptides (32);
- SDS–PAGE of peptides which have been linked to amino acid polymers (33).

Moreover, methods are being developed which avoid the use of radioisotopes, exploiting different properties such as the fluorescence of peptides which have been labelled with fluorescamine (34, 35), or separation of the phospho- and dephospho-peptides by capillary zone electrophoresis (36).

5. Assay of protein phosphatases using phosphorylated peptide substrates

Compared with protein kinases (especially PSKs), protein phosphatases do not appear to display a marked specificity for the primary sequence around their phosphoamino acid substrates (37). It is possible, however, to obtain a number of phosphorylated peptides which, irrespective of their relatedness to natural phosphorylation sites, are nevertheless readily and efficiently dephosphorylated by protein phosphatases, either serine/threonine or tyrosine specific. These peptides provide a valuable tool for monitoring these enzymes and are frequently exploited, especially in the case of protein–tyrosine phosphatases. A number of such peptides are listed in *Table 4*, with indications about the phosphatases that can be assayed with each of them. It should be noted once again that protein phosphatases are in general much more promiscuous than protein kinases with respect to their peptide substrates.

Table 4. Phosphopeptides used for the assay of protein phosphatases[a]

Phosphopeptide	Phosphorylating kinase[b]	Phosphatase(s) that can be assayed[c]
RRA**Tp**VA	PKA	PP2A and PP2C
RRREEE**Tp**EEE	CK2	PP2A
DLDVPIPGRFDRRV**Sp**VAAE	PKA	PP2B > PP2A >PP2C
ING**Sp**PR**Tp**PRRGQNR	Cyclin-dep.PK	PP2A (trimeric)
SpEEEEE	CK2	Acid/alkaline phosphatases
KKKKKRF**Sp**FKK**Sp**FKLSSF**Sp**FKKNKK	PKC	PP1, PP2A, PP2B, PP2C
EDNE**Yp**TA	Syk	PTPases
NIDGEVN**Yp**EE	c-Fgr	PTPases
AFLEDFFTSTEPQ**Yp**QPGENL	Csk	PTPases
TAEPD**Yp**GALYE		H PTPases β, LAR, CD45

[a] Data from ref. 37.
[b] Abbreviations: CK2, protein kinase CK2 or casein kinase 2; Csk, C-terminal Src protein kinase.
[c] Abbreviations: PP, protein phosphatase; PTPase, protein tyrosine phosphatase.

A complication in the case of phosphopeptides as substrates for protein phosphatases is the fact that they must be phosphorylated prior to use. This is normally carried out using [γ-^{32}P]ATP as phosphate donor and a suitable protein kinase as the catalyst. The latter will be chosen for its ability to phosphorylate the desired peptide, which may give rise to some problems. It should be noted in this respect that the best peptide substrates for protein phosphatases-2A and -2C contain phosphothreonine rather than phosphoserine. Since threonine is phosphorylated much less readily than serine by most kinases, the preparation of large amounts of phosphothreonyl peptides is sometimes troublesome. On the other hand, it should be remembered that it may not be necessary to achieve exhaustive phosphorylation of the peptide, nor to free the phosphopeptide of its non-phosphorylated form. It has been shown that even the presence of a large excess of the dephosphorylated peptide does not appreciably affect the kinetic parameters of most protein phosphatases, either serine/threonine or tyrosine specific. It is therefore common to use the mixture of the two forms without separation, ignoring the presence of the dephosphorylated peptide, and assuming that the phosphopeptide concentration corresponds to that of [^{32}P]phosphate incorporated into it.

5.1 Use of radiolabelled peptides for protein phosphatase assays

5.1.1 Preparation

[^{32}P]Phosphopeptides to be used as protein phosphatase substrates can be obtained by incubating the peptides with the suitable protein kinases, as described in Section 4.1. When very pure kinases are used, the incubation time can be prolonged up to some hours, in order to obtain as high a phosphorylation degree as possible. In some cases, the addition of a protease inhibitor cocktail to the reaction mixture can help in preventing peptide degradation during very long incubations. Particular care should be taken in the planning of the phosphopeptide preparation, avoiding the presence of any possible inhibitors of the protein phosphatase, since the phosphopeptide purification method might not allow their elimination.

Phosphorylation reactions are stopped by addition of 30% acetic acid, and separation of phosphopeptides from [γ-^{32}P]ATP is achieved, whenever applicable, by using the AG 1-X8 chromatography technique. The specific radioactivity of [γ-^{32}P]ATP needs to be quite high (2000–5000 cpm/pmol) and must be exactly known, since the concentration of the phosphopeptide will be assumed to be equal to that of the [^{32}P]phosphate in the peptide. After AG 1-X8 anion-exchange purification, the phosphopeptide is lyophilized and resuspended in the desired amount of water, adjusting the pH if necessary.

Protocol 7. Radioactive phosphopeptide preparation for protein phosphatase assay

Equipment and materials

- disposable columns (e.g. Pasteur pipettes plugged with glass wool, and a suitable rack to hold them)
- AG 1-X8 anion-exchange resin [equilibrated in 30% (v/v) acetic acid]
- 30% (v/v) acetic acid

Method

1. Carry out the phosphorylation reaction as described under *Protocol 1*, in a total volume of 200 μl.[a]
2. Stop the reaction by addition of 85 μl of 100% (v/v) acetic acid plus 215 μl 30% (v/v) acetic acid, to reach a total volume of 0.5 ml, and a final concentration of 30% (v/v) acetic acid.
3. Prepare a number of columns, corresponding to the number of samples, by pouring about 0.5 ml of AG 1-X8 resin into each column. Wash the resin extensively with 30% (v/v) acetic acid.
4. Load the samples on to the columns and collect the flow through fraction.
5. Wash three times with 1 ml of 30% (v/v) acetic acid and collect the eluate in the same tubes as the flow through, to a total volume of 3.5 ml.
6. Count a 20 μl aliquot out of the 3.5 ml by scintillation counting.[b]
7. Lyophilize the phosphopeptides eluted from the AG 1-X8 columns.[c]
8. Resuspend the lyophilized phosphopeptide in water, or buffer at the desired pH.
9. Check the concentration of the [^{32}P]phosphopeptide by counting 2 μl in a scintillation counter.[d]

[a] The incubation volume should be chosen dependending of the amount of phosphopeptide required. The amount of 100% and 30% acetic acid to be added at step 2 will change as a consequence.
[b] The result will give an estimate of the degree of phosphorylation, and will indicate the appropriate volume in which the phosphopeptide needs to be resuspended after lyophilization in order to obtain the desired concentration.
[c] Add an equal volume of water to the sample in 30% acetic acid, to allow it to freeze at –80°C.
[d] Divide the cpm in the phosphopeptide by the specific radioactivity of the ATP (cpm/pmol) to obtain the pmol of phosphopeptide contained in 2 μl, and hence calculate the phosphopeptide concentration.

5.1.2 Dephosphorylation reaction and assay

Phosphopeptide dephosphorylation reactions are performed under the optimal conditions for the particular protein phosphatase to be tested. Phosphopep-

tides are usually used at concentrations in the micromolar and submicromolar range for serine/threonine phosphatases and tyrosine phosphatase, respectively. Incubations (20–50 μl) are stopped after the appropriate time (5–30 min) by addition of trichloroacetic acid. The degree of dephosphorylation is assessed by quantification of the amount of [^{32}P]phosphate released, after its conversion into a phosphomolybdate complex. The complex can be extracted into an organic phase that is counted for ^{32}P radioactivity, while the radiolabelled phosphopeptide remains in the aqueous phase.

Protocol 8. Peptide phosphatase assay

Materials

- 5% (w/v) ammonium molybdate (prepared as in *Protocol 3*)
- TCA [trichloroacetic acid, 10% (w/v)]
- vortex mixer or rotational shaker
- water-saturated isobutanol/toluene (1:1) (prepared as in *Protocol 3*)
- liquid scintillation counter

Method

1. Stop the dephosphorylation reaction by adding 1.5 ml of 10% TCA plus 0.5 ml of 5% ammonium molybdate.
2. Add 2.5 ml of isobutanol/toluene (1:1) solution. Cap the tubes and shake for about 30 sec.[a]
3. Wait for the physical separation of the two phases; then withdraw 2 ml from the upper organic phase and count the radioactivity in 2 ml of scintillant.
4. Calculate the total amount of phosphate released by multiplying the cpm detected in the 2 ml by the factor 1.25.

[a] Mixing of the aqueous and organic phases can be obtained by vortexing, or on a rotational shaker.

5.2 Alternative methods for assaying protein phosphatases with phosphopeptides

Phosphopeptides can also be made by chemical synthesis (38, 39), the main advantage being the very high yield (approaching 100% phosphorylation) and the possibility to overcome the severe limitations imposed by the specificity and efficiency of the available protein kinases. The use of synthetic phosphopeptides may allow the assay of protein phosphatase activity without employing ^{32}P, whose short half-life discourages the expensive synthesis of radioactive phosphopeptides that can be used for only a limited period of time. This requires the development of reliable and sensitive techniques for the non-radioactive assay of phosphate release. Many attempts to achieve this have

been described, exploiting different detection methods. It is worth mention the following possibilities:

- reaction of phosphate with the reagent malachite green (40, 41), a colorimetric assay which allows the detection of nanomole amounts of phosphate;
- changes in ultraviolet absorption and fluorescence resulting from the dephosphorylation of a phosphotyrosyl peptide, which can be followed continuously (42);
- changes in fluorescence intensity in response to dephosphorylation of a properly designed phosphoseryl peptide, with a tryptophan residue adjacent to the phosphorylated site (43);
- decrease in fluorescent signal from an anti-phosphotyrosine antibody, due to the release of phosphate from the tyrosyl-phosphorylated substrate (44);
- the effect of unlabelled phosphopeptides as competitive inhibitors for the dephosphorylation of a standard radiolabelled substrate (39).

Acknowledgements

The authors are indebted to Giuseppe Tasinato for his invaluable help in designing and performing technical devices to improve the assay of peptide phosphorylation, and to our colleagues Arianna Donella Deana, Flavio Meggio, and Paolo Ruzza for providing unpublished data. The work done in the laboratory of the authors was supported by grants from the European Commission (BioMed-2), Harmenise Harvard Foundation, C.N.R. (Target Project on Biotechnology), Associazione Italiana Ricerca sul Cancro, M.U.R.S.T., and Italian Ministry of Health (Project AIDS).

References

1. Pinna, L.A., Ruzzene, M. (1996). *Biochim. Biophys. Acta*, **1314**, 19.
2. Ruzzene, M., Songyang, Z., Marin, O., Donella-Deana, A., Brunati, A.M., Guerra, B., Agostinis, P., Cantley, L.C., and Pinna, L.A. (1997). *Eur. J. Biochem.*, **246**, 433.
3. Gibbs, C.S. and Zoller, M.J. (1991). *Biochemistry*, **30**, 5329.
4. Gibbs, C.S. and Zoller, M.J. (1991). *J. Biol. Chem.*, **266**, 8923.
5. Sarno, S., Vaglio, P., Issinger, O.G., and Pinna, L.A. (1996). *J. Biol. Chem.*, **271**, 10595.
6. Sarno, S., Vaglio, P., Marin, O., Issinger, O.-G., Ruffato, K., and Pinna, L.A. (1997). *Biochemistry*, **36**, 11717.
7. Taylor, S.S., Radio-Andzelm, E., and Hunter, T. (1995). *FASEB J.*, **9**, 1255.
8. Hubbard, S.R. (1997). *EMBO J.*, **16**, 5572.
9. Leighton, I A , Dalby, L N , Caudwell, F B , Cohen, P T W , and Cohen, P (1995) *FEBS Lett* , **375**, 289

10. Songyang, Z., Blechner, S., Hoagland, N., Hoeckstra, M.F., Piwnica-Worms, H., and Cantley, L.C. (1994). *Curr. Biol.*, **4**, 973.
11. Songyang, Z., Carraway, K.L., III, Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., Lorenzo, M.J., Ponder, B.A.J., Mayer, B.J., and Cantley, L.C. (1995). *Nature*, **373**, 536.
12. Brunati, A.M., Donella-Deana, A., Ruzzene, M., Marin, O., and Pinna, L.A. (1995). *FEBS Lett.*, **367**, 149.
13. Lee, T.R., Niu, J., and Lawrence, D.S. (1995). *J. Biol. Chem.*, **270**, 27022.
14. Yan, X., Corbin, J.D., Francis, S.H., and Lawrence, D.S. (1996). *J. Biol. Chem.*, **271**, 1845.
15. Glass, D.B., Masaracchia, R.A., Feramisco, J.R., and Kemp, B.E. (1978). *Anal. Biochem.*, **87**, 566.
16. Marin, O., Meggio, F., and Pinna, L.A. (1994). *Biochem. Biophys. Res. Commun.*, **198**, 898.
17. Lasa-Benito, M., Marin, O., Meggio, F., and Pinna, L.A. (1996). *FEBS Lett.*, **382**, 149.
18. Marin, O., Meggio, F., Sarno, S., Andretta, M., and Pinna, L.A. (1994). *Eur. J. Biochem.*, **223**, 647.
19. Wilson, L.K., Dhillon, N., Thorner, J., and Martin, G.S. (1997). *J. Biol. Chem.*, **272**, 12961.
20. Kemp, B.E., Bylund, D.B., Huang, T.S., and Krebs, E.G. (1975). *Proc. Natl. Acad. Sci. USA*, **72**, 3448.
21. Meggio, F., Donella, A., and Pinna, L.A. (1976). *Anal. Biochem.*, **71**, 583.
22. Martin, J.B., and Doty, D.M. (1949). Anal. Chem., 21, 965.
23. Donella Deana, A., Meggio, F., and Pinna, L.A. (1979). *Biochem. J.*, **179**, 693.
24. Bylund, D.B., and Huang, T.S. (1976). *Anal. Biochem.*, **73**, 477.
25. Perich, J.W., Meggio, F., Reynolds, E.C., Marin, O., and Pinna, L.A. (1992). *Biochemistry*, **31**, 5893.
26. Boyle, W.J., Van der Geer, P., and Hunter, T. (1991). In *Methods in enzymology* (ed. T. Hunter and B.M. Sefton), Vol. 201, p. 111. Academic Press, London.
27. Hunter, T., Ling, N., and Cooper, J.A. (1984). *Nature*, **311**, 480.
28. Fiol, C.J., Wang, A., Roeske, R.W., and Roach, P.J. (1990). *J. Biol. Chem.*, **265**, 6061.
29. Budde, R.J.A., McMurray, J.S., and Tinker, D.A. (1992). *Anal. Biochem.*, **200**, 347.
30. Toomik, R., Ekman, P., Eller, M., Järv, J., Zaitsev, D., Myasoedov, N., Ragnarsson, U., and Engström, L. (1993). *Anal. Biochem.*, **209**, 348.
31. Goueli, B.S., Hsiao, K., Tereba, A., and Goueli, S.A. (1995). *Anal. Biochem.*, **225**, 10.
32. Lou, q., Wu, J., and Lam, K.S. (1996). *Anal. Biochem.*, **235**, 107.
33. Kameshita, I., and Fujisawa, H. (1996). *Anal. Biochem.*, **237**, 198.
34. Lutz, M.P., Pinon, D.I., and Miller, L.J. (1994). *Anal. Biochem.*, **220**, 268.
35. Seethala, R., and Menzel, R. (1997). *Anal. Biochem.*, **253**, 210.
36. Dawson, J.F., Boland, M.P., and Holmes, C.F.B. (1994). *Anal. Biochem.*, **220**, 340.
37. Pinna, L.A., Donella Deana, A. (1994). *Biochim. Biophys. Acta*, **1222**, 415.
38. Kitas, E.A., Knorr, R., Trzeciak, A., and bannwarth, W. (1991). *Helv. Chim. Acta*, **74**, 1315.
39. Perich, J.W., Ruzzene, M., Pinna, L.A., and Reynolds, E.C. (1994). *Int. J. Peptide Protein Res.*, **43**, 39.

40. Lanzetta, P.A., Alvarez, L.J., Reinach, P.S., and Candia, O.A. (1979). *Anal. Biochem.*, **100**, 95.
41. Martin, B., Pallen, C.J., Wang, J.H., and Graves, D.J. (1985). *J. Biol. Chem.*, **260**, 14932.
42. Zhang, Z.-Y., Maclean, D., Thieme-Sefler, A., Roeske, R.W., and Dixon, J.E. (1993). *Anal. Biochem.*, **211**, 7.
43. Wright, D.J., Noiman, E.S., Chock, P.B., and Chau, V. (1981). *Proc. Natl. Acad. Sci. USA*, **78**, 6048.
44. Babcook, J., Watts, J., Aebersold, R., and Ziltener, H.J. (1991). *Anal. Biochem.*, **196**, 245.
45. Kemp, B.E., Graves, D.J., Benjamini, E., and Krebs, E.G. (1977). *J. Biol. Chem.*, **252**, 4888.
46. Feramisco, J.R., Glass, D.B., and Krebs, E.G. (1980). *J. Biol. Chem.*, **255**, 4240.
47. Mitchell, R.D., Glass, D.B., Wong, C-W, Angelos, K.L., and Walsh, D.A. (1995). *Biochemistry*, **34**, 528.
48. Leader, D.P., Donella Deana, A. Marchiori, F., Purves, F.C., and Pinna, L.A. (1991). *Biochim. Biophys. Acta*, **1091**, 426.
49. Ferrari, S., Marchiori, F., Borin, G., and Pinna, L.A. (1985). *FEBS Lett.*, **184**, 72.
50. Proud, C.G., Colthrust, D.R., Ferrari, S., and Pinna, L.A. (1991). *Eur. J. Biochem.*, **195**, 771.
51. Yasuda, I., Kishimoto, A., Tanaka, S., Tominaga, N., Sakurai, A., and Nishizuka, Y. (1990). *Biochem. Biophys. Res. Commun.*, **166**, 1220.
52. Toomik, R., and Ek, P. (1997). *Biochem. J.*, **322**, 455.
53. Kielbassa, K., Müller, H-J., Meyer, H.E., Marks, F. and Gschwendt, M. (1995). *J. Biol. Chem.*, **270**, 6156.
54. Nishikawa, K., Toker, A., Johannes, F.J., Songyang, Z., and Cantley, L.C. (1997). *J. Biol. Chem.*, **272**, 952.
55. Alessi, D.R., Candwell, F.B., Andjelkovic, M., Hemmings, B.A., and Cohen, P. (1996). *FEBS Lett.*, **399**, 333.
56. Dale, S., Wilson, W.A., Edelman, A.M., and Hardie, D.G. (1995). *FEBS Lett.*, **361**, 191.
57. Mitchell, B.J., Stapleton, D., Mitchelhill, K.I., House, C.M., Katsis, F., Witters, L.A., and Kemp, B.E. (1996). *J. Biol. Chem.*, **271**, 28445.
58. Lee, J.C., Kwon, Y-G., Lawrence, D.S., and Edelman, A.M. (1994). *Proc. Natl. Acad. Sci. USA*, 91, 6413.
59. White, R.R., Kwon, Y.G., Taing, M., Lawrence, D.S., and Edelman, A.M. (1998). *J. Biol. Chem.*, **273**, 3166.
60. Pearson, R.B., Misconi, L.Y., and Kemp, B.E. (1986). *J. Biol. Chem.*, **261**, 25.
61. Heierhost, J., Tang, X., Lei, J., Probst, W.C., Weiss, K.R., Kemp, B.E., and Benian, G.M. (1996). *Eur. J. Biochem.*, **242**, 454.
62. Miyano, O., Kameshita, I., and Fujisawa, H. (1992). *J. Biol. Chem.*, **267**, 1198.
63. Stokoe, D., Caudwell, B., Cohen, P.T.W., and Cohen, P. (1993). *Biochem. J.*, **296**, 843.
64. Pearson, R.B., Mitchelhill, K.I., and Kemp, B.E. (1993). In *Protein phosphorylation: a practical approach* (ed. D.G. Hardie), p. 265. Oxford University Press, New York.
65. Zhang, J., Sanchez, R.J., Wang, S., Guarnaccia, C., Tossi, A., Zahariev, S., and Pongor, S. (1994). *Arch. Biochem. Biophys.*, **315**, 415.

66. Marin, O., Meggio, F., Daretta, G., and Pinna, L.A. (1992). *FEBS Lett.*, **301**, 111.
67. Srinivasan, J., Koszelak, M., Mendelow, M., Kwon, Y-G., and Lawrence, D.S. (1995). *Biochem. J.*, **309**, 927.
68. Garriga, J., Segura, E., Mayol, X., Grubmeyer, C., and Grana, X. (1996). *Biochem. J.*, **320**, 983.
69. Kameshita, I., Ishida, A., and Fujisawa, H. (1997). *J. Biochem.*, **122**, 168.
70. Gonzales, F.A., Raden, D.L., and Davis, R.J. (1991). *J. Biol. Chem.*, **266**, 22159.
71. Joseph, C.K., Buyn, H-S., Bittman, R., and Kolesnick, R.N. (1993). *J. Biol. Chem.*, **268**, 20002.
72. Okuno, S., Kitani, T., and Fujisawa, H. (1997). *J. Biochem.*, **122**, 337.
73. Mullinax, T.R., Stepp, L.R., Brown, J.R., and Reed, L.J. (1985). *Arch. Biochem. Biophys.*, **243**, 655.
74. Chan, K-F. J., Hurst, M.O., and Graves, D.J. (1982). *J. Biol. Chem.*, **257**, 3655.
75. Onorato, J.J., Palczewski, K., Regan, J.W., Caron, M.G., Lefkowitz, R.J., and Benovic, J.L. (1991). *Biochemistry*, **30**, 5118.
76. Purves, F.C., Donella-Deana, A., Marchiori, F., Leader, D.P., and Pinna, L.A. (1986). *Biochim. Biophys. Acta*, **889**, 208.
77. Wang, Q.M., Roach, P.J., and Fiol, C.J. (1994). *Anal. Biochem.*, **220**, 397.
78. Flotow, H., Graves, P.R., Wang, A, Fiol, C.J., Roeske, R.W., and Roach, P.J. (1990). *J. Biol. Chem.*, **265**, 14264.
79. Meggio, F., Perich, J.W., Reynolds, E.C., and Pinna, L.A. (1991). *FEBS Lett.*, **283**, 303.
80. Lam, K.S., Wu, J., and Lou, Q. (1995). *Int. J. Peptide Protein Res.*, **45**, 587.
81. Cheng, H.C., Nishio, H., Hatase, O., Ralph, S., and Wang, J.H. (1992). *J. Biol. Chem.*, **267**, 9248.
82. Marin, O., Donella-Deana, A., Brunati, A.M., Fischer, S., and Pinna, L.A. (1991). *J. Biol. Chem.*, **266**, 17798.
83. Edison, A.M., Barker, S.C., Kassel, D.B., Luther, M.A., and Knight, W.B. (1995). *J. Biol. Chem.*, **270**, 27112.
84. Ruzza, P., Calderan, A., Filippi, B., Biondi, B., Donella-Deana, A., Cesaro, L., Pinna, L.A., and Borin, G. (1995). *Int. J. Peptide Protein Res.*, **45**, 529.
85. Garcia, P., Shoelson, S.E., George, S.T., Hinds, D.A., Goldberg, A.R., and Miller, W.T. (1993). *J. Biol. Chem.*, **268**, 25146.
86. Bin, X., Bird, V.G., and Miller, W.T. (1995). *J. Biol. Chem.*, **270**, 29825.

tley

sub-
Vatl.

hos-

992)
iobi-
eins.

EXBIT 6

12

Analysis of Protein Kinase Substrate Specificity by the Use of Peptide Libraries on Cellulose Paper (SPOT-Method)

Werner J. Tegge and Ronald Frank

1. Introduction

Protein phosphorylation by protein kinases is the most important regulatory mechanism of cell function and signal transduction. In general, protein kinases exhibit specificities that are often primarily determined by the amino acids around the phosphorylation sites *(1)*. Identification of amino acids that contribute to substrate motifs are essential for the understanding of signal transduction pathways and for the development of specific peptide substrates and inhibitors. Many investigations with large numbers of individual peptides have been conducted in order to find high-affinity substrates as well as inhibitors *(2)*. Peptide libraries offer the possibility to investigate the sequence dependence of the phosphorylation more thoroughly and systematically and may even allow the *a priori* delineation of peptide substrates of uncharacterized protein kinases.

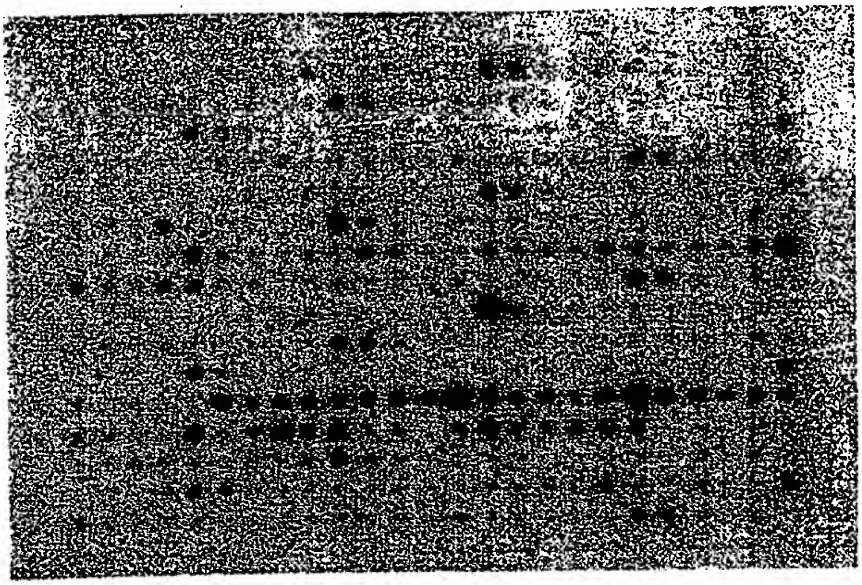
Recently, two new approaches have been described in this respect. Lam and coworkers have used one-bead one-peptide libraries of immobilized pentapeptides and heptapeptides comprising millions of individual sequences *(3)* generated by the method of "split synthesis" *(4)* (for a detailed description of the method *see* Chapter 10 in this volume). Songyang, Cantley, and coworkers have described an approach that uses a soluble library of 15mer peptides containing 8 degenerate positions adjacent to serine or tyrosine to evaluate substrate motifs of several serine/threonine and tyrosine kinases *(5,6)* (Chapter 11 in this volume).

We have developed a new method for the systematic investigation of the sequence-dependent specificity of protein kinases with peptide libraries on cel-

From: *Methods in Molecular Biology, vol. 87: Combinatorial Peptide Library Protocols*
Edited by: S. Cabilly © Humana Press Inc., Totowa, NJ

lulose paper, which is based on the SPOT-method *(7)* described in detail in a previous volume of this series *(8)* (*see also* Chapter 4). The method allows the identification of the major determining amino acids of substrate motifs and the systematic evaluation of every position. cAMP- and cGMP-dependent protein kinases (hereafter termed PKA and PKG, respectively) have been used as model enzymes for an evaluation of our method *(9)*. PKA has a well-defined substrate recognition motif. Thus, the literature data could serve as a measure of the performance of our approach and as a valuable guide during the development of the assay conditions. Both kinases are closely related to each other and share many similar features. While PKA displays a well-defined specificity with the general motif RRXS(A)X for substrate or inhibitory peptides *(10)*, PKG seems to have a less well-defined recognition sequence *(11,12)*.

In our investigations we used an iterative approach of constructing peptide libraries arranged in arrays of sublibraries that contain two defined amino acids each, similar to the mimotope strategy by Geysen et al. *(13)*. The best amino acid combination from a particular array was used throughout in the next one. The first generation had the general structure Ac-XXX12XXX, where X represents positions with equal distributions of all 20 natural amino acids (cysteine was used in its acetamidomethyl-protected form). The whole array represents a library of all 2.56×10^{10} possible octamers and each sublibrary with defined amino acids at positions 1 and 2 consists of 6.4×10^{7} sequences. Incubation with PKA gave a phosphorylation pattern with basic amino acids at positions 1 and 2 (**Fig. 1**). The sublibrary containing arginine at both positions shows particularly high activity. This is in agreement with results from soluble peptides, in which the general motif RXXRR/KXSX has been identified for this kinase. The two adjacent amino acids arginine-arginine in the middle seems to be the most strongly determining part of the whole motif. The second array had the general structure Ac-XXXRR12X. In this case the sequences with serine or threonine at position 2 gave the highest signals (**Fig. 2**). The discrimination of the enzyme at position 1 is much weaker. By applying this strategy, we have evaluated every position of the octameric sequence with both kinases and extended the sequences to decameric peptides. Promising peptides have been resynthesized in larger amounts by standard solid-phase peptide synthesis and enzyme kinetics have been determined. The decameric sequences that we obtained for PKG showed substrate specificites that were better than the ones known at that time. Exchange of serine for alanine resulted in a PKG-inhibitor with high affinity and specificity (W. Tegge and W. R. G. Dostmann, unpublished results). It should be kept in mind, though, that the approach presented here selects for substrates with a high V_{max} rather than for a low K_m, which is desired for good inhibitors. Further N- and C-terminal extensions of the length of the peptides to 14mers did not improve the substrates over decamers.



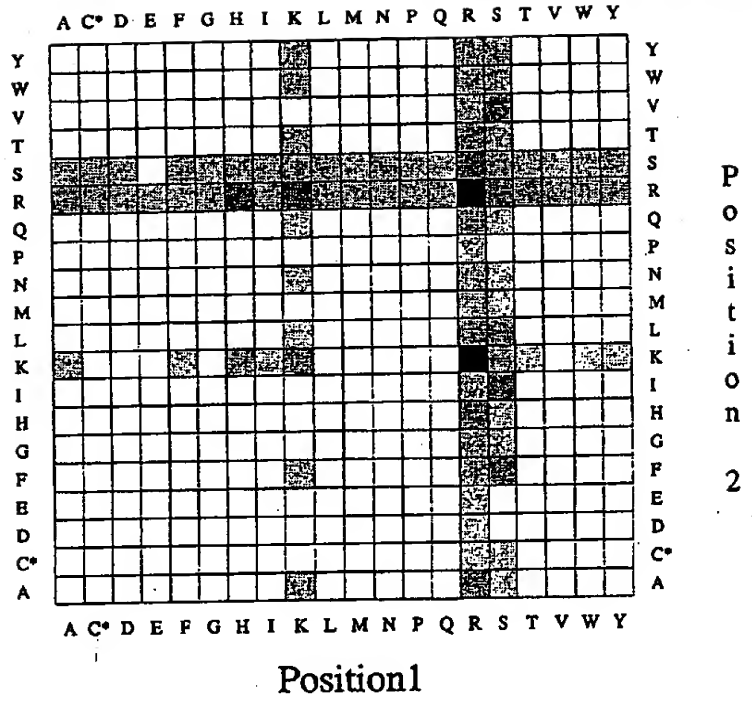


Fig. 1. Upper: PhosphorImager scan of the paper with the array Ac-XXX12XXX after phosphorylation by PKA. Four hundred sublibraries are arranged in a format of 16 × 25. The rows are arranged in an order so that 20 consecutive spots have a particular amino acid at position 1, and position 2 is one of the 20 amino acids, both positions being varied in alphabetical order (according to the single letter code). Lower: Quantified pattern of the 400 spots from the upper array, generated with the "Spectral" option of the program Corel Chart. The shading of each square corresponds with a linear dependence to the amount of phosphorylation of the corresponding spot on the paper. C* = Cys(Acm).

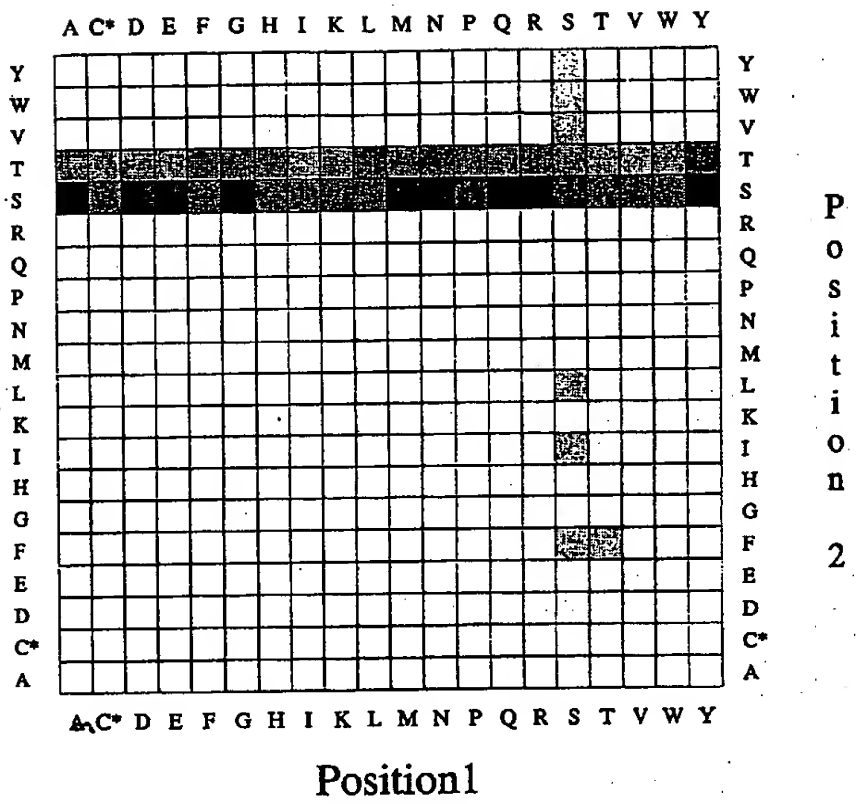


Fig. 2. Quantified pattern of the phosphorylation of the array Ac-XXXRR12X by PKA.

The iterative strategy described above is only an example. Of course, other procedures can be used for the delineation of the peptide motif. For example, it may be advantageous to predefine the position and the identity (e.g., serine or tyrosine) of the amino acid to be phosphorylated. With libraries of the structure XnS and SXn and/or XnY and YXn, the relative importance of N- and C-terminal positions to the substrate recognition and the question of whether a yet uncharacterized protein kinase belongs to a family of serine/threonine or tyrosine kinases may be evaluated. For a detailed discussion of such strategies *see* the chapter by Frank in this series *(8)* and references cited therein.

It should be considered that the SPOT procedure has a certain amount of scattering (up to 30%). In cases in which small differences between sublibraries of interest have been found, a reevaluation with an array presenting these sequences severalfold seems advisable before descisions about the structure of the next sublibrary array are being made. It should also be kept in mind that the amount of radioactivity incorporated into a particular sublibrary may be the result of the summation of a high number of sequences with rather low specificities rather than of a few sequences with very high activity, which is nor-

mally assumed for the planning of the next array. Furthermore, the significance of a residue at a certain position is influenced by the amino acids around it. These aspects imply that the final "optimal sequence" that has been found by such an interative approach may depend on the search strategy employed and may not necessarily be the best possible one. It can be expected, though, that it contains at least the major contributors to a motif.

Our SPOT approach has the advantage that every amino acid at every position of the sequence can easily be evaluated and compared without special technical considerations. This is important if, for example, not the identification of the best substrate motif but rather the largest differences between substrate properties for related enzymes is the major goal. Also, if more than one phosphorylation site is present in a sequence, their individual function as a phosphate acceptor needs to be evaluated. For this purpose individual spots can easily be punched or cut out of the array and analyzed. Peptide quantities per spot are in the range of 4–5 nmol, which is sufficient for microsequencing and/or amino acid analysis.

In general, for a successful application of methods based on short linear peptides, the site of phosphorylation (or generally the chemical modification) and at least some of the major contributors for the recognition motif must be accommodated into the length of the peptide that is evaluated. It should also be kept in mind, that only linear determinants are being identified. Interactions through secondary and/or tertiary structures that seem to be important for some classes of enzymes cannot be investigated with short peptides.

In conclusion, the SPOT approach can be expected to be generally applicable to the elucidation of protein kinase specificity and to the investigation of other enzymatic transformations.

2. Materials

The following materials and procedures work very well with PKA and PKG. For other kinases certain modifications may be necessary. The quantities are given for an assay involving one SPOT paper of standard size (8 × 12 cm). With an array of 425 spots the paper contains approx 2 µmol of peptide.

1. SPOT papers with libraries or peptide arrays according to the desired strategy (for the generation of these papers *see* **ref. *8*** and Chapter 4).
2. Reaction troughs with a lid, maybe of plexiglass of adequate thickness for shielding against the β-radiation of ^{32}P. The inner dimensions should be slightly larger than the paper. A very useful device is the Beta Work Box from Amersham (Amersham Buchler, Braunschweig, Germany), outer dimensions 185 × 115 × 80 mm, Code RPN 1539. The box can be used during the incubation and washing procedures of the SPOT papers and for storing the radioactive sheets.

3. An incubation chamber adjusted to 30°C containing a rocker for agitating the paper in the reaction trough during the incubation with the kinase.
4. A solution of 10 m*M* ATP, 100 µL per assay are required. Dissolve ATP disodium salt hydrate at 5.7 mg/mL in H_2O. Freeze stock solutions at –20°C. The stock should not be used for more than about 4 wk.
5. A solution of activator of suitable concentration, if required (e.g., 10 m*M* cAMP or cGMP in H_2O, always prepare fresh, because cAMP and cGMP hydrolyze rapidly).
6. 250 mL Incubation buffer: 50 m*M* MOPS, 200 m*M* NaCl, 1 m*M* MgAcetate, 0.4 m*M* EGTA, 1 mg/mL bovine serum albumin (BSA), pH 6.9 (adjusted with 1*N* NaOH). Prepare fresh. Set 8 mL of the buffer aside.
7. [γ-^{32}P]ATP (6,000 Ci/mmol, Amersham).
8. 1 L of a 1*M* solution of NaCl.
9. 100 mL of a stripping solution: 4 *M* guanidine hydrochloride, 1% SDS, 0.5% β-mercaptoethanol; for the removal of background, if desired. Prepare fresh, dissolve with warming.
10. A sonication bath with adjustable temperature.
11. Access to a PhosphorImager or StormSystem (Molecular Dynamics, Krefeld, Germany), or an X-ray film of sufficient size in combination with a flatbed scanner for the generation of a radioactivity image.
12. The program ImageQuant (Molecular Dynamics) can be used for the quantification of radioactivity. The option "Spectral" of the program Corel Chart (part of the Corel Draw program package) can be used for a convenient graphical presentation of the data (*see* **Fig. 1** and **2**).

3. Methods

1. Place the dry paper with the peptide array to be investigated into the incubation trough, moisten the paper with a few mL of ethanol, wash twice each with 50 mL incubation buffer and keep overnight in 100 mL of this buffer at room temperature (see **Note 1**).
2. Decant the buffer, add the 8 mL buffer that were set aside (*see* **Subheading 2., step 6**) and preincubate the paper at 30°C for a sufficient time (see **Note 2**).
3. Add 100 µL of 10 m*M* ATP and 100 µL each of the appropriate stock solutions of additional activators, if required (e.g., cAMP or cGMP) (*see* **Note 3**).
4. Add 10–100 µCi [γ-^{32}P]-ATP with the appropriate safety precautions for working with strong β-radiation.
5. Start the reaction by adding the kinase. Distribute the enzyme in the incubation buffer quickly and thoroughly. Final enzyme concentrations in our assays were 12.5 n*M* and 4 n*M* for PKA and PKG, respectively.
6. Incubate the mixture for 10 min at 30°C with slight agitation.
7. Decant the buffer solution and wash the paper 10 times with 100 mL each of 1*M* NaCl.
8. Wash several times with H_2O, add 100 mL of the stripping solution, and sonicate the paper for 1 h at 40°C to decrease the background level (*see* **Note 4**).

9. Wash the paper several times with water and ethanol, and then dry it (*see* **Note 5**).
10. Determine the radioactivity with the PhosphorImager or Storm system. Exposure times of the screen depend on the amount of radioactivity used and incorporated. Several h to 1 d was usually sufficient in our investigations. If no such system is available, use an X-ray film and scan the film after development.
11. Quantify the spots with the program ImageQuant by integrating uniformly sized circular areas that are positioned in the centers of the spots. Transfer the data ("sum above background," with background set to "0") to the program Corel Chart for a graphical presentation.

4. Notes

1. The ethanol treatment assures a good solubilization of the peptides in the aqueous buffer. If this step is not carried out and the buffer is added to the dry paper, the peptide spots appear as white areas on the paper and it may take several h or even d until they are completely solubilized. Treatment with the incubation buffer overnight blocks the paper that might otherwise adsorb the kinase. If the blocking is carried out for several days, it should be done at 5°C to prevent the growth of microorganisms, unless the buffer contains preservative.
2. The wet paper contains approx 2 mL of buffer. Addition of another 8 mL brings the volume to 10 mL. A preincubation period of at least 1 h should be used. If the trough has been kept at 5°C, extend that period of time. The trough has a fairly high heat capacity so that it takes a considerable time to warm it up to 30°C.
3. If the catalytic subunit of PKA is being used, no cAMP or any other activator needs to be added. PKG contains the regulatory unit as part of the primary structure and requires the addition of cGMP at a final concentration of 100 µ*M*.
4. This step may not be necessary if fresh and clean buffers and enzymes are being used. If small radioactive spots appear on or between the peptide spots after scanning the paper, include **step 8** into the procedure.
5. Drying can be carried out by leaving the paper uncovered for about 1 h, or if the process shall be speeded up, by using an electrical hair dryer.

References

1. Kemp, B. E. and Pearson, R. B. (1990) Protein kinase recognition sequences motifs. *Trends Biochem. Sci.* **15**, 342–346.
2. Pearson, R. B. and Kemp, B. E. (1990) Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulation. *Methods Enzymol.* **200**, 62–81.
3. Wu, J., Ma, Q. N., and Lam, K. S. (1994) Identifying substrate motifs of protein kinases by a random library approach. *Biochemistry* **33**, 14,825–14,833.
4. Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. (1991) General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493.

5. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H., and Cantley, L. C. (1994) Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr. Biol.* **4,** 973–981.
6. Songyang, Z. (1995) Catalytic specificity of protein tyrosine kinases is critical for selective signalling. *Nature* **373,** 536–539.
7. Frank, R. (1992) Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* **48,** 9217–9232.
8. Frank, R. and Overwin, H. (1996) SPOT-synthesis: epitope analysis with arrays of synthetic peptides prepared on cellulose membranes in, *Methods in Molecular Biology,* vol. 66: *Epitope Mapping Protocols* (Morris, G. E., ed.), Humana, Totowa, NJ, pp. 149–169.
9. Tegge, W., Frank, R., Hofmann, F., and Dostmann, W. R. G. (1995) Determination of cyclic nucleotide-dependent protein kinase substrate specificity by the use of peptide libraries on cellulose paper. *Biochemistry* **34,** 10,569–10,577.
10. Glass, D. B., Cheng, H.-C., Mende-Müller, L., Reed, J., and Walsh, D. A. (1989) Primary structural determinants essential for potent inhibition of cAMP-dependent protein kinase by inhibitory peptides corresponding to the active portion of the heat-stable inhibitor protein. *J. Biol. Chem.* **264,** 8802–8810.
11. Butt, E., Abel, K., Krieger, M., Palm, D., Hoppe, V., Hoppe, J., and Walter, U. (1994) cAMP- and cGMP-dependent protein kinase phosphorylation sites of the focal adhesion vasodilator-stimulated phosphoprotein (VASP) in vitro and in intact human platelets. *J. Biol. Chem.* **269,** 14,509–14,517.
12. Glass, D. B. (1990) Substrate specificity of the cyclic GMP-dependant protein kinase in, *Peptides and Protein Phosphorylation* (Kemp, B. E., ed.), CRC, Boca Raton, FL, pp. 209–238.
13. Geysen, H. M. and Mason, T. J. (1993) Screening chemically synthesized peptide libraries for biologically-relevant molecules. *Bioorg. Med. Chem. Lett.* **3,** 397–404.

EXHIBIT 7

11

The Use of Peptide Library for the Determination of Kinase Peptide Substrates

Zhou Songyang and Lewis C. Cantley

1. Introduction

Protein phosphorylation plays a crucial role in regulating a plethora of intracellular biological activities. Because protein phosphorylation is a reversible reaction, it allows living cells to reset to the basal state rapidly after stimulation. There are hundreds of protein kinases involved in this general signaling machinery. These kinases are classified into three different categories based on their abilities to phosphorylate serine/threonine, or tyrosine, or both residues. A comparison of primary sequences of protein kinases has indicated that the catalytic site (the kinase domain) is highly conserved, suggesting a common ancestor for these kinases *(1)*. However, different kinases have evolved to function distinctly in response to diverse cellular stimuli. The specificities of protein kinases has thus become a critical issue in understanding signal transduction.

The specificity of a protein kinase is governed by a number of factors, including the intracellular localization of the kinase and its substrates. The most important factor is the specificity of the kinase domain. Studies of protein Ser/Thr kinases, including cAMP-dependent protein kinase (PKA), indicate that the kinase domain recognizes specific primary sequences around the phopshorylation site *(2,3)*.

The specificities of a few protein Ser/Thr kinases have been verified by amino acid substitutions on the basis of known in vivo substrates. However, this conventional approach has several drawbacks. First, it is extremely expensive and time-consuming to synthesize and assay all the possible substitutions. Since 9–12 amino acids of the substrate peptide are likely to contact the active site cleft of a kinase *(3)*, there are approx 20^{10} or 10^{13} distinct peptides to

From: *Methods in Molecular Biology, vol. 87: Combinatorial Peptide Library Protocols*
Edited by: S. Cabilly © Humana Press Inc., Totowa, NJ

consider. Second, this method is less applicable to cases in which the in vivo targets of protein kinases have not yet been identified. To overcome these problems, several new strategies based on combinatorial chemistry have been used. One approach takes advantage of a technique that simultaneously synthesizes large numbers of degenerate peptides on a solid matrix *(4,5*, and Chapters 10 and 12*)*. The specificity of the assayed kinase can be deduced by decoding (e.g., sequencing) immobilized peptides. However, this approach may be problematic because substrate phosphorylation on solid matrices may be different from that in solution. Therefore, peptide substrates identified from these methods may not be physiologically relevant.

We have developed an oriented peptide library technique to rapidly determine optimal sequences for protein kinases based on a strategy similar to that used for Src-homology 2 (SH2) domains *(6,7)*. The kinase of interest is added to a soluble mixture of billions of distinct peptides, each of identical length and orientation, with only a single amino acid capable of being phosphorylated located in the middle. The small fraction of phosphorylated peptides is quantitatively separated from the bulk of nonphosphorylated peptides and the mixture is sequenced. A comparison of the abundance of amino acids at each degenerate position surrounding the phosphorylation site to the abundance at the same position in the starting mixture indicates preferred amino acids at each position. This technique not only predicts an optimal sequence from a single experiment without prior knowledge of in vivo phosphorylation sites, but also provides information about the relative importance of each position for selectivity and which amino acids are tolerated. In this chapter, we will focus on this oriented peptide library technique.

2. Materials

2.1. Chemicals

2.1.1. Standard Reagents for Peptide Synthesis and Sequencing

See Chapters 1, 3, 4, 8 and 13 for details.

2.1.2. Protein Kinase Assay

1. Kinase buffer for Ser/Thr kinases: 50 m*M* Tris-HCl, 10 m*M* $MgCl_2$, 1 m*M* dithiothreitol (DTT).
2. Kinase buffer for Tyr kinases: 50 m*M* Tris-HCl, 10 m*M* $MnCl_2$, 1 m*M* DTT.
3. ATP and [γ-^{32}P]-ATP.

2.1.3. DEAE Column

1. DEAE sephacel (Sigma, St. Louis, MO).
2. 30% Acetic acid.

2.1.4. Ferric Column

1. Iminodiacetic acid (IDA)-coupled agarose beads (Pierce, Rockford, IL).
2. 20 m*M* Ferric chloride.
3. Buffer A: 50 m*M* MES, 1 *M* NaCl, pH 5.5.
4. Buffer B: 2 m*M* MES, pH 6.0, or distilled H_2O.
5. Buffer C: 500 m*M* NH_4HCO_3, pH 8.0.
6. 100 m*M* EDTA, pH 8.0.

2.2. Equipment and Supplies

1. Peptide synthesizer (ABI 431A) and sequencer (ABI 477A).
2. Peristaltic pumps.
3. Polypropylene columns or syringes.
4. Speed-Vac.

3. Methods

The method shown here has been used successfully in determining substrate specificities of many protein kinases. Similar to our earlier approach in determining motifs for SH2 domains *(6)*, this procedure quantitatively separates the degenerate phosphopeptide products from the bulk of nonphosphorylated peptides.

3.1. Design of Degenerate Peptide Libraries

Peptide libraries for protein kinases could be classified by the number of fixed residues. For instance, primary libraries are those that only use amino acids capable of being phosphorylated (Tyr, Ser, or Thr) to fix and orient the libraries. Secondary libraries fix additional residues that are preferred by some kinases (e.g., Ser–Pro for cyclin-dependent kinases). The most important criteria in designing the peptide libraries for protein kinases are the following:

1. Proper orientation of the peptide library: for protein kinases, this is easily achieved because the amino acids capable of being phosphorylated (Tyr, Ser, and Thr) can be used to orient the library. For example, the following peptide library (primary library, Ser Degenerate Library) was constructed for serine/threonine kinases: Met-Ala-Xxx-Xxx-Xxx-Xxx-Ser-Xxx-Xxx-Xxx-Xxx-Ala-Lys-Lys-Lys, where Xxx indicates all amino acids except Trp, Cys, Tyr, Ser, or Thr *(7)*. Amino acids capable of being phosphorylated were omitted at all degenerate positions to ensure that the only potential site of phosphorylation was the Ser at residue 7. Each degenerate position was thus fixed relative to the Ser and the phosphorylated peptides were in phase when sequenced. To design a secondary library, a second residue in addition to the amino acid capable of being phosphorylated is fixed. For example, this secondary library (Ser–Pro library) was constructed for cyclin-dependent kinases: Met-Ala-Xaa-Xaa-Xaa-Xaa-Ser-Pro-

Xaa-Xaa-Xaa-Ala-Lys-Lys-Lys, where Xaa indicates all amino acids except Trp and Cys *(7)*.

2. Selection of amino acids at the degenerate positions: As discussed above, amino acids capable of being phosphorylated (Tyr, Ser, and Thr) are usually avoided in the degenerate positions for primary libraries. However, for secondary libraries, Tyr, Ser, and Thr can be included in the degenerate positions because phosphorylation of these residues at the degenerate position is negligible. In the case of Ser–Pro library, the chance of having Ser and Pro adjacent to each other at the degenerate positions is quite small ($<5 \times 18^{-2}$, ~1.5%). Thus, peptides phosphorylated at positions other than the fixed one would not interfere with the sequencing of oriented peptides. For all peptide libraries, Trp and Cys have been omitted to avoid problems with sequencing and oxidation. These two residues could be substituted into specific locations once the optimal peptides have been determined. If 15 different amino acids are present in any one of the eight degenerate positions (for primary libraries), the total theoretical degeneracy of the library is $15^8 = 2{,}562{,}890{,}625$. A primary peptide library for tyrosine kinases (Tyr-Kinase Substrate Library) was constructed in this fashion by except that residue 7 was Tyr *(8)*.
3. Other general considerations: The length of the peptide libraries can vary, but the number of degenerate positions should not exceed 15. A library of 15 degenerate positions already has 20^{15} different molecules. We start with a library containing eight degenerate positions because four residues N-terminal and four C-terminal to the phosphorylation site. This is the region most likely to be involved in catalytic recognition based on the motifs that had been determined for protein kinases. Placing a short leading sequence before the degenerate positions is generally beneficial. Taking the Ser Degenerate Library (Met-Ala-Xxx-Xxx-Xxx-Xxx-Ser-Xxx-Xxx-Xxx-Xxx-Ala-Lys-Lys-Lys), for example, the Met-Ala sequence at the N-terminus of the peptide libraries provides two amino acids to verify that peptides from this mixture are being sequenced. Sequencing of these two residues also allows quantification. Similarly, the Ala at residue 12 makes it possible to quantify and estimate how much peptide has been lost during sequencing. The poly-Lys tail prevents wash-out during sequencing and improves the solubility of the mixture (no solubility problems occurred at neutral pH and 5 mg/mL concentration).

3.2. Peptide Library Synthesis

Synthesis of the degenerate peptide libraries is accomplished according to the standard BOP/HOBt coupling protocols using a Peptide BioSynthesizer (ABI 431A).

1. At the degenerate positions, add equal moles of 15 different Fmoc-blocked amino acids simultaneously at a 10-fold excess to the coupling resin. The ratio of input Fmoc-blocked amino acids sometimes needs to be adjusted on different synthesizers to achieve an even distribution of degenerate amino acids.

2. Deprotect and cleave the resins by trifluroacetic acid (TFA).
3. Sequence the peptide libraries to confirm that all amino acids are present at similar amounts (within a factor of 3) at all degenerate positions.

3.3. Protein Kinases: Preparation and Kinase Assays

Protein kinases used in this study can be obtained through different sources (*see* **Note 1**). Kinase reactions can be performed with soluble or immobilized kinases.

1. Add the protein kinase to 300 µL of solution containing 1 mg of degenerate peptide mixture, 100 µ*M* ATP with a trace of [γ-^{32}P]-ATP (roughly 6×10^5 cpm) in kinase buffer.
2. Incubate at 25–30°C for 2 h to phosphorylate roughly 1% of the peptide mixture.
3. Terminate the reaction with the addition of acetic acid to a final concentration of 15%.

3.4. Phosphopeptide Separation

3.4.1. Purification of Peptide Libraries on DEAE Column

After an incubation period at 25°C, the peptide supernatant is removed and diluted with 300 µL of 30% acetic acid.

1. Add the mixture to a 1-mL DEAE-sephacel column previously equilibrated with 30% acetic acid.
2. Elute the column with 30% acetic acid (*9*). Discard the first 600 µL flow through and collect the next 1 mL (*see* **Note 2**).
3. Lyophilize the collected fraction on a Speed-Vac.

3.4.2. Ferric Chelation Column

A ferric chelation column (IDA beads) is used for separation of phosphopeptides (*see* **Note 3**). This column has been used in the past to separate tryptic phosphopeptides of phosphorylated proteins from the bulk of nonphosphorylated tryptic peptides (*10,11*). However, we discovered that in order to accomplish a quantitative removal of the nonphosphorylated peptides from the phosphopeptides without loss of a subfraction of phosphopeptides, it is necessary to change the loading and elution conditions from published procedures. A typical running profile is shown in **Fig. 1**.

1. Charge a ~1-mL column of IDA beads with 5 mL of 20 m*M* ferric chloride at 0.5 mL/min.
2. Wash with 8 mL of water at 1 mL/min.
3. Wash with 6 mL of buffer C at 1 mL/min.
4. Wash again with 6 mL of water.

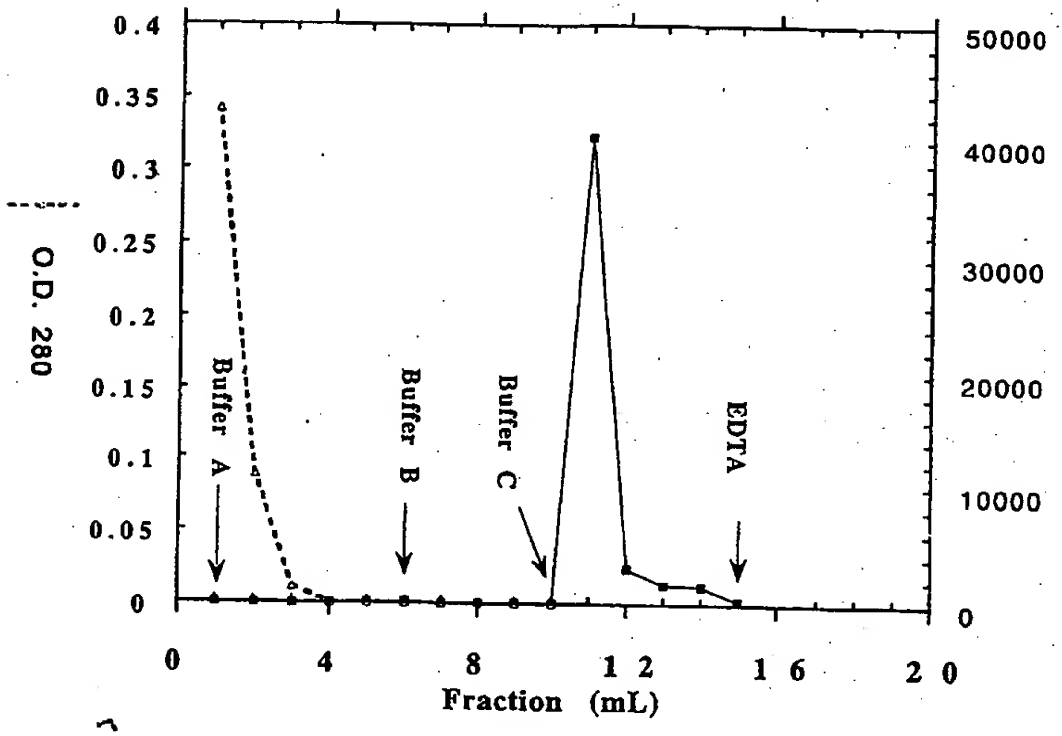


Fig. 1. Quantitative separation of nonphosphorylated and phosphopeptides on a ferric-chelating column. Approximately 1% of the peptides in the Tyr-Kinase Substrate Library were phosphorylated by polyoma middle T/pp60c-src. After separating the [γ-^{32}P]-ATP from the peptides on a DEAE-Sephacel column. the peptide mixture was loaded on a column of ferric-IDA beads (*see* Methods). The column was eluted with 5 mL of buffer A, 5 mL of buffer B, 4 mL of buffer C, and 2 mL of 100 m*M* EDTA (pH 8.0). All elutions were at 0.5 mL/min and 0.5-mL fractions were collected. The amount of peptide at each fraction was estimated by absorbency at UV 280 nm (fractions 1–10) and phosphopeptide was estimated by radioactivity. Less than 0.1% of the total nonphosphorylated peptides eluted at fractions 11–14 as judged by Tyr at cycle 7 in the sequence of this mixture. Greater than 90% of the radioactivity applied eluted in fractions 11–14.

5. Equilibrate with 6 mL of buffer A.
6. Dissolve the dried sample of peptide/phosphopeptide mixture in 200 µL of buffer A and load onto the ferric column.
7. Wash the column with 5 mL of buffer A followed by 5 mL of buffer B or H_2O at 0.5 mL/min.
8. Elute the phosphopeptides with 4 mL of buffer C.
9. Elute Fe^{+3} with 100 m*M* EDTA, pH 8.0.
10. Collect the buffer C eluate, which contains phosphopeptide, and lyophilize several times to get rid of most of the ammonium bicarbonate salt. Resuspend the phosphopeptide mixture in water, adjust to neutral pH, and sequence.

3.5. Sequencing and Data Analysis

Sequencing of the phosphopeptide mixture reveals the abundance of amino acids at each degenerate position. It is also necessary to sequence the original peptide library mixture. To determine the optimal peptide motif for a protein kinase, in theory, the abundance of each amino acid at a given cycle in the sequence of the phosphopeptide mixture could be divided by the abundance of the same amino acid at the same cycle of the starting mixture. In this way, variations in the abundance of amino acids at a particular residue (i.e., residue 3, the first degenerate residue) in the starting mixture or variations in yield of amino acids in the sequencer are divided out. If the kinase is insensitive to the amino acid at residue 3 (i.e., four residues N-terminal of the phosphoserine), then the relative abundance of all amino acids at this cycle in the phosphopeptide mixture will be the same as in the starting mixture, and all bars in the graph will have equal height.

In some experiments in which only approximately 0.5% of the total mixture is phosphorylated, a correction should be made for the ~0.1% of the nonphosphorylated peptides that are eluted with buffer C. The contamination with nonphosphorylated peptides can be estimated from the quantity of Ser (or Tyr) at cycle 7 since residues do not show up if phosphorylated. Control experiments are conducted in which the peptides are subjected to a mock phosphorylation. The same column protocol is used and the fractions in which phosphopeptides usually elute are collected and sequenced. These fractions are usually rich in Asp and Glu at every degenerate cycle, presumably owing to interaction of these residues with the Fe^{3+}. The amino acid abundance at each cycle from this control is subtracted from the kinase experiment to correct for the background. To calculate the relative preference of amino acids at each degenerate position, the corrected data are then compared to the starting mixture to create the ratios of abundance of amino acids. The sum of the abundance of each amino acid at a given cycle is normalized to 15 or 18 (the number of amino acids present) so that each amino acid would have a value of 1 in the absence of selectivity at a particular position. In theory, any value greater than 1 should indicate preference for the corresponding amino acid. However, because of the complexity of the data and calculation, we found that values higher than 1.5 are generally reliable. The whole process is summarized as following:

1. Normalize the amount of each amino acid at the degenerate positions: we routinely normalize the total amount of amino acids (in picomoles) of a degenerate position to that of the first degenerate position.
   a. $P(ij)$ indicates amount for amino acid j at position i for the kinase experiment.

b. *Pn*(ij) indicated normalized *P*(ij).
c. *Pn*(ij) = *P*(ij) × Sum(*P*1)/Sum(*P*i);

2. Normalize sequences of the control experiment and original peptide library as in **step 1**.
   a. *cn*(ij) indicates the normalized amount for amino acid j at position i in the control experiment.
   b. *Rn*(ij) indicates the normalized amount for amino acid j at position i in the original peptide library.
3. Subtract the control experiment values from the kinase experiment: *Pn*(ij) - *K* × *Cn*(ij); *K* can be calculated by the relative amount of fixed Ser or Tyr, *K* = *P*(ser)/*C*(ser).
4. Calculate the relative abundance *A*(ij) = [*Pn*(ij) - *K* × *Cn*(ij)]/*Rn*(ij).
5. Normalize to the total number of amino acids included at the degenerate position: *An*(ij) = *A*(ij) × 18/Sum[*A*ij]. *An*(i,j) represents the enrichment value for amino acid j at position i.

After the calculation, graphic plots showing enrichment values of amino acids at all degenerate positions can be generated to reveal the substrate preference for individual protein kinases. In **Fig. 2**, the specificity of PKA determined with the primary Ser degenerate peptide library (MAXXXXSXX-XXAKKK) was plotted using the Kaleidagraph program. Ultimately, confidence in this procedure is provided by the reproducibility of the results obtained with a given kinase and by consistency of predicted optimal substrates with known substrates of the same protein kinases.

3.6. Using the Predicted Optimal Peptide Substrates to Study Signal Transduction by Protein Kinases

3.6.1. Predicting Kinase Substrates

The optimal substrates deduced from this method are extremely useful in predicting the in vivo targets of various protein kinases. In particular, this

Fig. 2. *(see opposite page)* Substrate specificity of cAMP-dependent protein kinase detected by the degenerate peptide library. The phosphopeptide produced by phosphorylating the Ser-Kinase Substrate Library with PKA were sequenced. Each box indicates the relative abundance of the 15 amino acids at a given cycle of sequencing. For example, box A is cycle 3, the first degenerate position in the library mixture. Cycle 7 (not shown) is the site of phosphorylation (phosphoserine). Therefore, boxes A, B, C, and D indicate amino acid preferences at –4, –3, –2, and –1 N-terminal of the phosphorylation site and boxes E, F, G, and H indicate preferences at +1, +2, +3, and +4 C-terminal of the phosphorylation site. The columns represent average values from two independent experiments. The bars indicate the differences between the two experiments. Abbreviations for amino acid residues are: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; Y, Tyr.

as in

in the

in the

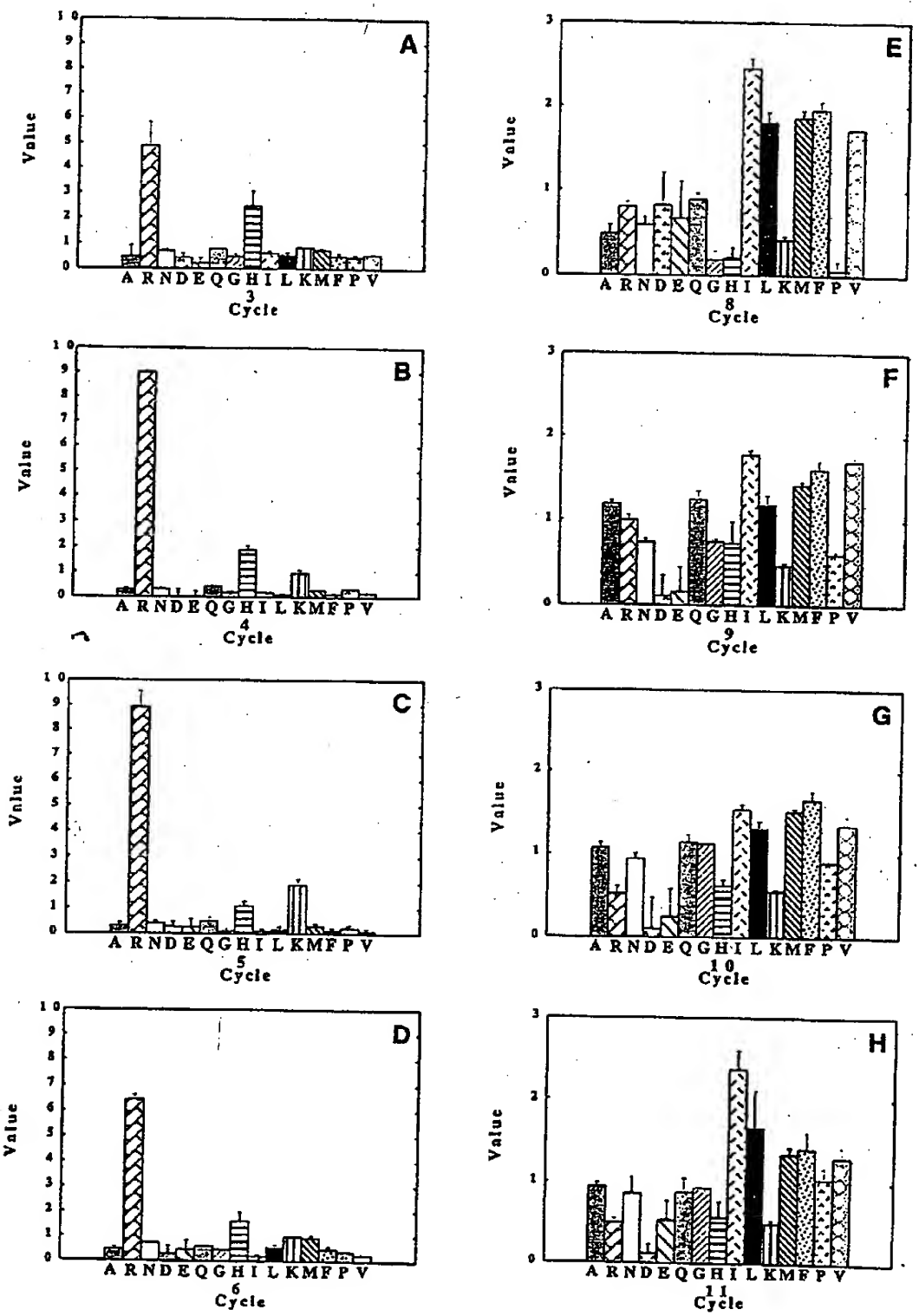
Cn(ij);
er).

posi-
ie for

mino
trate
PKA
SXX-
onfi-
ained
with

ul in
. this

inase
spho-
indi-
g. For
cle 7
B, C,
phos-
id +4
from
two
u; F,
g; S,



method can rapidly identify the optimal peptide substrates of a kinase without previous knowledge of the kinase. One can take these optimal peptide sequences to search protein databases using Blast, Fasta, or Findpatterns in the Genbank GCG program (GCG). Matched proteins are likely in vivo substrates of the kinase studied. Meanwhile, one can also scan the sequence of a protein to see whether it contains any potential phosphorylation sites for a given protein kinase. Both approaches provide a shortcut in understanding signaling pathways regulated by the protein kinases.

3.6.2. Developing Inhibitors of Protein Kinases

The predicted optimal peptide substrates can facilitate the design of inhibitors of protein kinases. First, peptide or peptide mimetic inhibitors can be made based on the optimal peptides. Second, the optimal peptides can be used as specific probes to screen chemical libraries for potential inhibitors. Moreover, structural analysis of kinases complexed with their optimal peptides would also assure a basis for modeling and designing drugs that specifically intervene in protein kinase-mediated signaling.

4. Notes

1. Most commonly, the kinases of interest are overexpressed as recombinant proteins in bacteria and eukaryotic cells. For most kinases, expression in insect cells (sf9 cells) using baculoviruses is desirable and often yields active enzymes. Expressed protein kinases can be purified by conventional liquid chromatography (e.g., FPLC), affinity chromatography, or simple immunoprecipitation. The amount of enzyme required to phosphorylate enough peptides for sequencing varies and depends on the specific activity of individual kinases. In general, we use microgram quantities of kinases in our experiments.
2. Under these conditions, peptide mixtures are in the void volume because of their poly-Lys tail, while ATP and denatured protein kinases are retained on the column. In initial experiments the fractions from the column were analyzed for peptide, phosphopeptide, [γ-^{32}P]-ATP, and $^{32}PO_4$ by phosphocellulose (P81) paper, TLC, or SDS PAGE. It was determined that after the first 600 µL void volume, the next 1 mL contained both phosphorylated and nonphosphorylated peptides but was free of [γ-^{32}P]-ATP. Since the peptide fraction is free of [γ-^{32}P]-ATP, the radioactivity in this fraction provides an initial estimate of the fraction of the total peptide mixture that was phosphorylated.
3. Although the ferric chelation column efficiently separates phosphorylated peptides from the unphosphorylated species, a small percentage (~0.1%) of degenerate unphosphorylated peptides rich in acidic amino acids are also copurified because these peptides can bind weakly to the ferric column. This could be a problem for peptide libraries for which acidic residues (Asp and Glu) are fixed. In order to lower the background, acidic residues should be avoided at the fixed positions of peptide libraries. If acidic amino acids must be included at the

fixed positions, one could increase the pH of buffer A to 6.5 and reduce the volume of IDA beads.

It is important to purify enough phosphopeptides (e.g., up to 1% input peptide library) for sequencing such that phosphorylated peptides are in great excess of the contaminating unphosphorylated peptides. In a reaction in which 1% of the peptide mixture is phosphorylated, the total quantity of phosphopeptides is (1.8 m*M*) × (0.3 mL) × (0.01) = 5.4 nmol. Typically, about 1–2 nmol of phosphopeptide mixture is added to the sequencer. This means that in a cycle in which all 15 residues are equally abundant, the yield of each amino acid is (1 nmol) × (1/15) = 60 pmol.

Acknowledgment

We thank Dan Liu, Kermit L. Carraway, and Anhco Nguyen for critical reading of this manuscript.

References

1. Hanks, S. K., Quinn, A. M., and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241,** 42–52.
2. Glass, D. B., Cheng, H. C., Mueller, L. M., Reed, J., and Walsh, D. A. (1989) Primary structural determinants essential for potent inhibition of cAMP-dependent protein kinase by inhibitory peptides corresponding to the active portion of the heat stable inhibitor protein. *J. Biol. Chem.* **264,** 8802–8810.
3. Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253,** 414–420.
4. Geysen, H. M., Meloen, R. H., and Barteling, S. J. (1984) Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. USA* **81,** 3998–4002.
5. Wu, J., Ma, Q., N., and Lam, K. S. (1994) Identifying substrate motifs of protein kinases by a random library approach. *Biochemistry* **33,** 14,825–14,833.
6. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) SH2 domains recognize specific phosphopeptide sequences. *Cell* **72,** 767–778.
7. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica, W. H., and Cantley, L. C. (1994) Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr. Biol.* **4,** 973–982.
8. Songyang, Z., Carraway III, K. L., Eck, M. J., Harrison, S. C., Feldman, R. A., Mohammadi, M., Schlessinger, J., Hubbard, S. R., Smith, D. P., Eng, E., Lorenzo, M. J., Ponder, B. A. J., Mayer, B. J., and Cantley, L. C. (1995) Catalytic specificity of protein-tyrosine kinases is critical for selective signaling. *Nature* **373,** 536–539.

9. Kemp, B. E., Benjamini, E., and Krebs, E. G. (1976) Synthetic hexapeptide substrates and inhibitors of 3':5'-cyclic AMP-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **73,** 1038–1042.
10. Muszynska, G., Andersson, L., and Porath, J. (1986) Selective adsorption of phosphoproteins on gel-immobilized ferric chelate. *Biochemistry* **25,** 6850–6853.
11. Muszynska, G., Dobrowolska, G., Medin, A., Ekman, P., and Porath, J. O. (1992) Model studies on iron(III) ion affinity chromatography. II. Interaction of immobilized iron(III) ions with phosphorylated amino acids, peptides and proteins. *J. Chromatogr.* **604,** 19–28.

EXHIBIT 8

[6] Detection and Identification of Substrates for Protein Kinases: Use of Proteins and Synthetic Peptides

By LORENTZ ENGSTRÖM, PIA EKMAN, ELISABET HUMBLE, ULF RAGNARSSON, and ÖRJAN ZETTERQVIST

Two types of intracellular phosphoproteins are known: enzymes that are intermediately phosphorylated at their active sites and proteins that are phosphorylated by protein kinases on serine, threonine, or tyrosine residues. In both cases the proteins are phosphorylated after their synthesis.

Intermediate phosphoryl enzymes are formed rapidly, usually within milliseconds, on incubation of the enzymes with their substrates. At the active sites of phosphoglucomutase and alkaline phosphatase, a serine residue is phosphorylated, whereas in all other known cases a histidine, a glutamic acid, or an aspartic acid residue is the acceptor of the phosphoryl groups.[1]

When phosphoproteins are formed through the action of protein kinases, steady-state levels are usually reached only after several minutes. This type of phosphorylation mainly takes place on serine residues, while phosphorylation on threonine residues represents a minor fraction. Fairly recently, the phosphorylation of tyrosine residues by specific protein kinases has been described, both in normal cells and after infection with tumor virus. The amount of tyrosine-bound phosphate in normal cells is very low[2] compared with that of serine- and threonine-bound phosphate, which in several mammalian tissues is about 0.5–1.0 μmol per gram of wet tissue.[3,4]

Since the detection of a protein kinase in 1954,[5] several different protein kinases and their substrates have been found in a number of tissues and cell fractions. The enzymes differ with regard to substrate specificity and factors that influence their activity, as seen in the table. In the present chapter, the detection and use of substrates for protein kinases active on serine and threonine residues will be discussed, with emphasis on substrates and kinases present in mammalian tissues.

[1] M. Weller, *in* "Protein Phosphorylation," p. 163. Pion, London, 1979.
[2] T. Hunter and B. M. Sefton, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1311 (1980).
[3] T. A. Langan, *in* "Regulation of Nucleic Acid and Protein Biosynthesis," BBA Library, Vol. 10, p. 233. Elsevier, Amsterdam, 1967.
[4] H. Forsberg, Ö. Zetterqvist, and L. Engström, *Biochim. Biophys. Acta* **181**, 171 (1969).
[5] G. Burnett and E. P. Kennedy, *J. Biol. Chem.* **211**, 969 (1954).

ISBN 0-12-182007-6

Detection of Protein Phosphorylation in Crude Tissue Extracts

The search for a particular protein phosphorylation can be made with use of whole animals, perfused tissues, intact cells, cell fractions, or crude tissue extracts. Although with intact cells there may be a lower risk of artifactual phosphorylation, crude tissue extracts usually offer a more convenient system for an initial scanning. In this section, experiments on crude extracts will be described. Phosphorylation in intact cells is discussed in a later section.

The crude tissue extract may consist in isolated cell supernatant or extracts of the particulate cell fractions. When the extracts are incubated with [γ-^{32}P]ATP and Mg^{2+}, considerable incorporation of [^{32}P]phosphate into proteins occurs, owing to the presence of protein kinases and endogenous substrates. This has been illustrated in experiments with rat liver cell sap.[6] A fairly large incorporation of [^{32}P]phosphate, i.e., near 1 nmol per milligram of cell sap protein, has been observed after incubation with 5 m*M* [^{32}P]ATP for 60 min at 30°. By incubation of the same material for a short period at low temperature, e.g., 15 sec at 0°, the incorporation of [^{32}P]phosphate into intermediate phosphoryl enzymes can be estimated, as was demonstrated for rat liver cell sap. This incorporation is of a much lower degree than that due to the protein kinase reactions and amounts to about 0.04 nmol per milligram of cell sap protein.[6,7]

A significant part of the ^{32}P-labeling is accounted for by a few, relatively abundant proteins. However, the detection of phosphorylation of a particular, minor component may be of equal interest. This usually requires a highly efficient separation method, such as the two-dimensional polyacrylamide gel electrophoresis of O'Farrell.[8] In the few cases in which the phosphoprotein is identified and specific antibodies are available, immunoprecipitation followed by one-dimensional polyacrylamide gel electrophoresis in sodium dodecyl sulfate may be used.

To achieve detectable ^{32}P-labeling of as many phosphoproteins of the extract as possible, several aspects have to be considered.

1. Although the apparent K_m with respect to ATP is rather low (10^{-5} to 10^{-4} *M*) for a number of protein kinases, the concentration of [^{32}P]ATP should initially be in the millimolar range to allow extensive ^{32}P labeling before the [^{32}P]ATP is hydrolyzed by ATP hydrolases.

2. To detect phosphorylation of a minor protein, the specific radioactivity of the [^{32}P]ATP should be high enough to give a detectable spot on

[6] O. Ljungström and L. Engström, *Biochim. Biophys. Acta* **336**, 140 (1974).
[7] Ö. Zetterqvist, *Biochim. Biophys. Acta* **141**, 533 (1967).
[8] P. H. O'Farrell, *J. Biol. Chem.* **250**, 4007 (1975).

an autoradiogram of a polyacrylamide gel electrophoresis. With enforcing screens and preexposure of the film with a photo flash, about 50 dpm of ^{32}P per square centimeter is detectable in 24 hr.[9] A specific radioactivity of 20 dpm/pmol would then permit the detection of a compound representing 1/1000 of the ^{32}P-labeled phosphoproteins of cell sap from 1 mg of liver tissue (i.e., from about 50 μg of cell sap protein).

3. Allowance should be made for the endogenous ATP present in the extract, if this compound has not been removed on a Sephadex G-50 column.

4. The specific radioactivity of the [^{32}P]ATP will decrease during the incubation, owing to the action of ATP hydrolases and adenylate kinase. To permit true estimates of the phosphorylation, [^{32}P]ATP therefore has to be isolated and the specific radioactivity of the γ-phosphorus determined.[6]

5. Phosphorylated or phosphorylatable sites may be more sensitive to proteolysis than the rest of the protein.[10] Inhibitors of proteolytic enzymes should therefore be added to the incubation mixture to reduce this risk.

6. Phosphorylatable proteins may have been partially phosphorylated in the cell. Thus a small or slow incorporation of [^{32}P]phosphate might be due to the presence of bound, nonradioactive phosphate at the beginning of the incubation, and the observed ^{32}P labeling will rather be a consequence of the turnover of the phosphate caused by the activity of phosphoprotein phosphatases present in the system.

The method of interruption of the phosphorylation greatly influences the amount and type of [^{32}P]phosphoproteins obtained. When further studies on native phosphoprotein are to be made, the reaction may be interrupted by the removal of free Mg^{2+} with EDTA. Phosphoprotein phosphatases may be inhibited by sodium fluoride or orthophosphate.[11] The ^{32}P-labeled phosphoproteins are separated from most of the low molecular weight, ^{32}P-labeled compounds by rapid gel filtration. It is important to keep in mind, however, that considerable amounts of these compounds may be adsorbed to the protein of crude extracts after the gel filtration. To obtain reliable estimates of the true protein phosphorylation,

[9] R. A. Laskey and A. D. Mills, *FEBS Lett.* **82**, 314 (1977).

[10] G. Bergström, P. Ekman, E. Humble, and L. Engström, *Biochim. Biophys. Acta* **532**, 259 (1978).

[11] T. S. Ingebritsen and D. M. Gibson, *in* "Recently Discovered Systems of Enzyme Regulation by Reversible Phosphorylation" (P. Cohen, ed.), p. 63. Elsevier/North-Holland, Amsterdam, 1982.

the adsorbed material must be removed by denaturation procedures (see below).

To prove definitely that an apparent protein phosphorylation is due to covalent binding to serine and threonine residues, it is necessary to isolate the phosphoamino acid after degradation of the protein, generally by partial acid hydrolysis in 2 *M* HCl for 10–20 hr at 100°, a method originally introduced by Lipmann.[12] After lyophilization of the acid hydrolysates, the phosphoamino acids are generally isolated by chromatography[13] or electrophoresis.[14] In this context, the importance of using more than one system at the identification of a particular phosphoamino acid is emphasized.

Identification of Protein Kinase Activities

If phosphorylation of an unknown protein substrate has been found in a crude tissue extract, attempts to identify the protein kinase activity and the substrate can be performed in parallel. In some cases, the substrate protein and the protein kinase activity may be enriched in the same fraction during a preliminary purification. Thus, rat liver pyruvate kinase and cyclic AMP-dependent protein kinase copurified during an acid precipitation (pH 5.5) and an ammonium sulfate fractionation step, which facilitated the identification of liver pyruvate kinase as substrate of cyclic AMP-dependent protein kinase.[15]

Use of Activators and Inhibitors of Protein Kinases

Addition of different activators of protein kinases (see the table), e.g., cyclic AMP, cyclic GMP, calcium ions with or without calmodulin, or phosphatidylserine and diolein, to crude systems may increase the rate of the phosphorylation reaction and thus point to the type of protein kinase active in the system.

The cyclic AMP-dependent protein kinase activity of a crude fraction can be inhibited by adding the protein inhibitor described by Walsh.[16] Inhibition of calcium-stimulated protein kinases can be achieved by the

[12] F. Lipmann, *Biochem. Z.* **262**, 3 (1933).
[13] L. Engström, *Biochim. Biophys. Acta* **52**, 49 (1961).
[14] J. A. Cooper, N. A. Reiss, R. J. Schwartz, and T. Hunter, *Nature (London)* **302**, 218 (1983).
[15] O. Ljungström, G. Hjelmquist, and L. Engström, *Biochim. Biophys. Acta* **358**, 289 (1974).
[16] D. A. Walsh, C. D. Ashby, C. Gonzalez, D. Calkins, E. H. Fischer, and E. G. Krebs, *J. Biol. Chem.* **246**, 1977 (1971).

addition of EGTA. The concomitant inhibition of the phosphorylation of the unidentified protein will give an indication as to which protein kinase is active.

Addition of Purified Protein Kinases

If the phosphorylation of an endogenous protein is enhanced by the addition of a purified kinase preparation, this indicates that the corresponding protein kinase is responsible for the phosphorylation of that protein when studied in the crude extract. However, it should be kept in mind that many proteins are phosphorylated by more than one protein kinase, as has been most clearly demonstrated for rabbit muscle glycogen synthase (see the table). In such cases, it seems to be common that different residues are phosphorylated by the different protein kinases. One way to establish that different residues are involved is to compare the patterns of the [^{32}P]phosphopeptides obtained after degradation by proteolytic enzymes and by partial acid hydrolysis.[17]

Tissue Distribution of a Protein Kinase Active on an Identified Substrate

If the substrate of a protein kinase has been identified and is available for enzyme assays, while the protein kinase itself is unidentified, investigation of the tissue distribution of the protein kinase activity[18–21] may facilitate the latter identification. This approach was used to identify the calcium-activated, phospholipid-dependent protein kinase described by Nishizuka and collaborators[22] as one of the enzymes that phosphorylates human fibrinogen.[23]

Identification of Natural Substrates for Protein Kinases

Preliminary Characterization of a Phosphorylatable Component

Some guidance for the identification of phosphorylatable proteins may be obtained from the estimation of their native and subunit molecular

[17] C. Milstein and F. Sanger, *Biochem. J.* **79,** 456 (1961).
[18] J. F. Kuo and P. Greengard, *Proc. Natl. Acad. Sci. U.S.A.* **64,** 1349 (1969).
[19] J. F. Kuo, *Proc. Natl. Acad. Sci. U.S.A.* **71,** 4037 (1974).
[20] J. F. Kuo, R. G. G. Andersson, B. C. Wise, L. Mackerlova, I. Salomonsson, N. L. Brackett, N. Katoh, M. Shoji, and R. W. Wrenn, *Proc. Natl. Acad. Sci. U.S.A.* **77,** 7039 (1980).
[21] R. Minakuchi, Y. Takai, B. Yu, and Y. Nishizuka, *J. Biochem. (Tokyo)* **89,** 1651 (1981).
[22] Y. Takai, A. Kishimoto, Y. Iwasa, Y. Kawahara, T. Mori, and Y. Nishizuka, *J. Biol. Chem.* **254,** 3692 (1979).
[23] P. Papanikolaou, E. Humble, and L. Engström, *FEBS Lett.* **143,** 199 (1982).

weights, e.g., by gel chromatography and by polyacrylamide gel electrophoresis in sodium dodecyl sulfate under reducing conditions, respectively. Other properties may also be investigated, e.g., their behavior in ion-exchange chromatography in different systems or the isoelectric point as determined by isoelectric focusing or chromatofocusing.

The amount of the phosphorylatable component is also useful for its identification. This may be estimated from the maximal [^{32}P]phosphate incorporation obtained, under the assumption that only one phosphate group is incorporated per subunit.

If the phosphorylatable protein is an adaptive enzyme, some clues to its identity may be provided by feeding the experimental animals with different diets and then determining the amounts of phosphorylatable proteins. The tissue distribution and subcellar location of the phosphorylatable component may also be helpful in the identification.

Identification from Known Hormonal Effects

The effect of many hormones at the molecular level is to increase the intracellular concentration of cyclic AMP, whose main, or perhaps sole, effect in mammals is to activate cyclic AMP-dependent protein kinase. The physiological actions of these hormones are thus exerted through phosphorylation of specific proteins. An unknown substrate of cyclic AMP-dependent protein kinase in a tissue should thus be related to some process that is regulated by hormones acting on the tissue via cyclic AMP. The identification of liver pyruvate kinase type L as a substrate of cyclic AMP-dependent protein kinase was facilitated by this approach.[15] Glucagon is the main hormone that increases the concentration of cyclic AMP in the liver. An important physiological effect of glucagon is to raise the level of blood glucose. This is accomplished partly by increasing the mobilization of glycogen and partly by stimulating gluconeogenesis with an increased synthesis of phosphoenolpyruvate from pyruvate.[24] On this basis, a protein in rat liver cell sap that was found to be phosphorylated under the influence of cyclic AMP and had the same subunit molecular weight and approximately the same concentration as pyruvate kinase could be identified as this enzyme.

In a similar way, Avruch and collaborators have identified a protein of fat and liver tissues, whose phosphorylation is stimulated by insulin and is identical to ATP citrate-lyase.[25]

[24] G. A. Robison, R. W. Butcher, and E. W. Sutherland, "Cyclic AMP," p. 232. Academic Press, New York, 1971.

[25] M. C. Alexander, E. M. Kowaloff, L. A. Witters, D. T. Dennily, and J. Avruch, *J. Biol. Chem.* **254**, 8052 (1979).

Studies of Amino Acid Sequences as a Means of Identifying Protein Substrates

A number of substrates of cyclic AMP-dependent protein kinase, e.g., pyruvate kinase, glycogen synthase, and phosphorylase kinase show similarities between their phosphorylated sites (see the table). Concerning liver pyruvate kinase, it has been found that the native conformation of the enzyme is not a prerequisite for phosphorylation, since both alkali-inactivated enzyme and a cyanogen bromide fragment from the same protein are phosphorylated at a greater rate than the native enzyme, although the final degree of phosphorylation is the same.[26] These observations initiated phosphorylation experiments with synthetic peptides representing part of the phosphorylatable site of pyruvate kinase, in which the essential role of the amino acid sequence around the phosphorylatable amino acid residue was demonstrated. Thus, by studying known primary structures of proteins, it should be possible to find substrates of cyclic AMP-dependent protein kinase. This method led to the finding that fibrinogen is a substrate *in vitro* of cyclic AMP-dependent protein kinase.[27]

However, the mere existence of a favorable amino acid sequence in a protein does not make it a protein kinase substrate. Another prerequisite is that the phosphorylation site be accessible to the protein kinase. In fact, a protein may be made phosphorylatable to a certain extent by denaturation.[28] Obviously, caution is needed in the interpretation of phosphorylation *in vitro*. Eventually, its occurrence *in vivo* always has to be considered.

Synthetic Peptides as Substrates of Protein Kinases

[^{32}P]Phosphopeptides were derived from phosphorylated pig and rat liver pyruvate kinase[29,30] and shown to exhibit one distinct feature, namely, two vicinal arginine residues. This feature, together with results from experiments with denatured pyruvate kinase,[26] induced our group[31]

[26] E. Humble, L. Berglund, V. Titanji, O. Ljungström, B. Edlund, Ö. Zetterqvist, and L. Engström, *Biochem. Biophys. Res. Commun.* **66**, 614 (1975).
[27] L. Engström, B. Edlund, U. Ragnarsson, U. Dahlqvist-Edberg, and E. Humble, *Biochem. Biophys. Res. Commun.* **96**, 1503 (1980).
[28] D. B. Bylund and E. G. Krebs, *J. Biol. Chem.* **250**, 6355 (1975).
[29] G. Hjelmquist, J. Andersson, B. Edlund, and L. Engström, *Biochem. Biophys. Res. Commun.* **61**, 559 (1974).
[30] B. Edlund, J. Andersson, V. Titanji, U. Dahlqvist, P. Ekman, Ö. Zetterqvist, and L. Engström, *Biochem. Biophys. Res. Commun.* **67**, 1516 (1975).
[31] Ö. Zetterqvist, U. Ragnarsson, E. Humble, L. Berglund, and L. Engström, *Biochem. Biophys. Res. Commun.* **70**, 696 (1976).

and other investigators[32] to explore synthetic peptides based on the phosphorylatable site of liver pyruvate kinase.

Peptide Synthesis

Synthesis of peptides consisting of up to about 10 amino acid residues can conveniently be carried out by the solid-phase method.[33] The early and simple manual version[34] is quite satisfactory as long as only a moderate number of peptides are required, and it is very suitable for gaining some practical experience of solid-phase peptide synthesis.[35] If large series of peptides are required, however, the synthetic manipulations can be speeded up considerably by using an automatic synthesizer.

Synthesis of Leu-Arg-Arg-Ala-Ser-Val-Ala. As an example, the synthesis of a heptapeptide, representing the phosphorylatable site of rat liver pyruvate kinase, will be described in some detail.

Chloromethylated, 1% cross-linked polystyrene (Biobeads, Bio-Rad) is esterified with Boc-alanine.[34,35] This resin, 1.75 g, containing 0.400 mmol of Boc-Ala, is loaded into the reaction vessel of a Beckman Model 990 peptide synthesizer. The Boc group is removed by exposure to a solution of 33% trifluoroacetic acid in dichloromethane for 30 min. Free amino groups are liberated by neutralization with 10% triethylamine in the same solvent for a total of 10 min. After further careful washing with CH_2Cl_2, the next amino acid, valine, is attached as its Boc derivative (2.5 equivalents) in CH_2Cl_2 with the aid of dicyclohexylcarbodiimide (also 2.5 equivalents) for 2 hr. To ensure the highest possible yield in this step, the coupling is repeated once with the same amounts of fresh reagents before proceeding.

The protecting Boc group is removed from the valine residue as described above and the next amino acid, serine, is attached, using the conditions described above. The serine derivative used is Boc-Ser (Bzl). Similarly, Boc-Ala, Boc-Arg(NO_2) (twice), and Boc-Leu are added. As Boc-Arg(NO_2) does not dissolve in CH_2Cl_2, a mixture of dimethylformamide and CH_2Cl_2 (2 and 9 ml, respectively) is used in these coupling steps.

The peptide is cleaved from the resin with liquid hydrogen fluoride.[35] Simultaneously, the groups protecting the side chains are removed from the peptide. After extraction of the polymer with 10% acetic acid, the peptide solution is lyophilized.

[32] B. E. Kemp, D. J. Graves, E. Benjamini, and E. G. Krebs, *J. Biol. Chem.* **252**, 4888 (1977).
[33] R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149 (1963).
[34] R. B. Merrifield, *Biochemistry* **3**, 1385 (1964).
[35] J. M. Stewart and J. D. Young, "Solid Phase Peptide Synthesis." Freeman, San Francisco, California, 1969.

An aliquot (about 100 mg of the product) is subjected to chromatography on a 1.4 × 16 cm column of carboxymethylcellulose. After equilibration with an ammonium acetate buffer, pH 5.5, 0.025 *M* in ammonium ions, elution is performed with a 600-ml linear gradient of 0.025 to 0.25 *M* ammonium acetate, at a flow rate of about 20 ml/hr. The peptide, which appears in the middle of the gradient, is traced by its absorbance at 235 nm, appropriate fractions are lyophilized, the residue is dissolved in water, and the solution is again lyophilized to give about 50 mg of a crystalline material. This preparation initially contains insignificant amounts of ammonium ions as seen at amino acid analysis. A typical analysis gave Arg 2.03, Ser 0.96 (after correcting for 11% decomposition during the acid hydrolysis), Ala 2.00, Val 0.99, Leu 1.01, and a peptide content of 62%.

The purity of this and similar peptides can be further determined by HPLC. The basicity and hydrophilicity of the peptide require a specially designed ion-pair chromatographic system.[36] Incidentally, the same system can be applied to separate the peptide from the corresponding phosphopeptide and also to separate different phosphopeptides from each other.[37]

Minimum Peptide Substrates for Cyclic AMP-Dependent Protein Kinase

Smaller peptides, lacking N-terminal Leu, Leu-Arg, and Leu-Arg-Arg or C-terminal Ala and Val-Ala, can be similarly prepared and assayed as substrates for the cyclic AMP-dependent protein kinase. With use of a set of these peptides, the peptide Arg-Arg-Ala-Ser-Val was shown to represent the minimum substrate phosphorylated at a significant rate.[31]

By a similar approach, a second type of substrate of cyclic AMP-dependent protein kinase, based on the sequence of a fragment from the β subunit of phosphorylase kinase,[38] has been explored.[39] In this case, the minimum substrate proved to be Arg-Thr-Lys-Arg-Ser-Gly-Ser-Val (Ser at the seventh position being phosphorylated). Although the two arginine residues of this peptide were not located next to each other, both were apparently essential to the phosphorylation, as shown by the substitution of glycine for either arginine residue. In addition, substitution of glycine for the lysine residue showed that the lysine was also of significance to the phosphorylation, although the influence of this exchange on the apparent K_m was less than with the arginine residues. The second hydroxyamino

[36] B. Fransson, U. Ragnarsson, and Ö. Zetterqvist, *J. Chromatogr.* **240**, 165 (1982).
[37] B. Fransson, U. Ragnarsson, and Ö. Zetterqvist, *Anal. Biochem.* **126**, 174 (1982).
[38] S. J. Yeaman, P. Cohen, D. C. Watson, and G. H. Dixon, *Biochem. J.* **162**, 411 (1977).
[39] Ö. Zetterqvist and U. Ragnarsson, *FEBS Lett.* **139**, 287 (1982).

acid, threonine, on the other hand, seemed to be replaceable without significant changes of the kinetic parameters.

Cyclic GMP-Dependent Protein Kinase

The specificity of cyclic GMP-dependent protein kinase, obtained from pig lung,[40] has also been investigated with use of synthetic peptides.[41,42] This enzyme seems to have structural requirements that are similar, although not identical, to those of cyclic AMP-dependent protein kinase. One observed difference is that, for a significant rate of phosphorylation, the cyclic GMP-dependent protein kinase seems to require two amino acid residues C-terminal to the phosphorylatable serine, compared with only one in the two minimum peptides for cyclic AMP-dependent protein kinase.

Techniques for Quantification of Phosphorylation of Proteins and Synthetic Peptides

Proteins

Incubation Mixtures. The concentration of protein in crude extracts may amount to about 10–15 mg/ml. When purified protein substrate is used, a practical concentration is around 2 mg/ml, or 50 μM. A specific radioactivity of [^{32}P]ATP of about 20 dpm/pmol for the γ-phosphorus and a 5 mM concentration of free Mg^{2+} are generally adequate. Several types of buffers have been used,[43] although in some cases orthophosphate has been shown to inhibit protein kinase activity.[44] However, in crude extracts this inhibition may only be apparent, since orthophosphate inhibits phosphoprotein phosphatase and will retard the removal of unlabeled phosphate present in the protein *in vivo*.

For experiments with the Ca^{2+}, calmodulin-activated protein kinase, 50 μM Ca^{2+}, and 0.57 μM calmodulin (10 μg/ml) may be used. The cyclic nucleotide-dependent protein kinases need about 1 μM cAMP or cGMP in a system with purified substrate and protein kinase. This concentration should be increased to 100 μM if the incubation mixture is suspected to contain diesterase activity, which hydrolyzes the cyclic nucleotide. The

[40] K. Nakazawa and M. Sauo, *J. Biol. Chem.* **250**, 7415 (1975).
[41] T. M. Lincoln and J. D. Corbin, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3239 (1977).
[42] B. Edlund, Ö. Zetterqvist, U. Ragnarsson, and L. Engström, *Biochem. Biophys. Res. Commun.* **79**, 139 (1977).
[43] S. A. Rudolph and B. K. Krueger, *Adv. Cyclic Nucleotide Res.* **10**, 107 (1979).
[44] J. H. Wang, J. T. Stull, T.-S. Huang, and E. G. Krebs, *J. Biol. Chem.* **251**, 4521 (1976).

diesterase may also be inhibited by the addition of 1 m*M* methylxanthine. The Ca^{2+} concentration required for phospholipid-activated protein kinases is in the order of 10^{-7} to 10^{-4} *M*, and concentrations exceeding 0.5 m*M* are usually inhibitory. A commonly used concentration of phosphatidylserine is 10–25 μg/ml.

In systems that are not extensively purified, the phosphorylation should be preferably performed at 30° for a rather short time, i.e., 5–10 min, since contaminating proteases and phosphoprotein phosphatases may decrease the recovery of the protein-bound [^{32}P]phosphate. Experiments with pure systems may be continued for about 2 hr when the maximal incorporation is to be estimated.

Termination of the Reaction and Analysis of ^{32}P Labeling. Two methods of interrupting the reaction and determining the phosphorylation of the total protein of the sample will be described. In the first method, the incubation is interrupted by applying a 25-μl aliquot of the incubation mixture to a Whatman 3 MM paper disk, 11 × 11 mm. The paper is then immediately placed in a stainless, wire-netting support with separate compartments for each paper, which stands in a beaker containing ice-cold 10% trichloroacetic acid and 50 m*M* phosphoric acid. The presence of 50 m*M* phosphoric acid in the trichloroacetic acid is important for preventing the adsorption of a possible trace amount of [^{32}P]orthophosphate present in the [^{32}P]ATP preparation. Free [^{32}P]ATP is removed by washing the papers for 10 min in the acid under magnetic stirring. This washing is repeated twice with fresh acid. The paper disks are then dried in ethanol for 5 min and, finally, in ethyl ether for 5 min. The papers are left in the wire-netting support throughout the washing and drying procedure. This procedure is similar to that described by Corbin and Reimann.[45] However, with the wire-netting support the washing is more efficient and more reproducible. The dry filter paper disk is placed on the bottom of an empty scintillation vial, and the Čerenkov radiation is measured in a liquid scintillation counter.

When it is suspected that a protein is not quantitatively adsorbed to the paper (which is the case for instance with fructose-1,6-bisphosphatase) or when the incubation volume has to be increased to facilitate detection of the incorporated [^{32}P]phosphate, precipitation with trichloroacetic acid in a test tube is preferred. It should be noted, however, that the concentration of acid required for quantitative precipitation can vary between different proteins. In most cases, a final trichloroacetic acid concentration of 10% is sufficient. Some basic proteins, however, require

[45] J. D. Corbin and E. M. Reimann, this series, Vol. 38 [41], p. 287.

the addition of tungstate to be precipitated.[46] The trichloroacetic acid may be added as solid acid or, preferably, as an aliquot of a concentrated stock solution. If the protein concentration is below 1 mg/ml, a carrier protein, for instance bovine serum albumin, can be added just before the addition of trichloroacetic acid. After immersion of the mixture in an ice-water bath for 10 min, the precipitate is collected by centrifugation at 1000 *g* for 5 min, dissolved in 0.5 *M* NaOH, and immediately reprecipitated by trichloroacetic acid at the appropriate final concentration. The protein should be dissolved and reprecipitated totally 4 times. After the last centrifugation, the protein is dissolved in 0.5 *M* NaOH and the Čerenkov radiation is measured.[47] Serine- and threonine-bound phosphate is fairly easily removed from phosphoproteins in strong alkaline solutions by β-elimination, even at room temperature.[48] Therefore, any treatment of this type of phosphoprotein with strong alkaline solutions should be short and preferentially performed in an ice-water bath. On the other hand, the bound phosphate remains very stable when exposed to strong acid at room temperature. Treatment with 1 *M* HCl for 30 min at room temperature will therefore remove histidine- and aspartic acid-bound phosphate without losses of serine- or threonine-bound phosphate.[7]

Preparation of Native Phosphoprotein. When it is desirable that a purified protein should remain in a native state after phosphorylation, the phosphoprotein and ATP may be separated by Sephadex G-50 gel chromatography. The phosphorylation reaction can be stopped by the addition of EDTA before the chromatography if the reaction time is critical. Otherwise, the reaction will be stopped upon separation of protein from ATP on the column. The volume of the sample should not exceed 10% of the Sephadex G-50 column volume, in order to get a good separation. The use of a phosphate buffer with a concentration of at least 50 m*M* is of advantage to decrease the adsorption of [^{32}P]ATP and trace amounts of [^{32}P]orthophosphate that are usually present in the [^{32}P]ATP preparation. Rechromatography usually reduces the adsorption even further. The degree of adsorption may be determined by treating an aliquot as described in the preceding paragraph.

Synthetic Peptides

Incubation Mixtures. Synthetic peptides have been used as substrates of various protein kinases, such as cyclic AMP-dependent protein kinase,[31,32] cyclic GMP-dependent protein kinase,[41,42] phosphorylase ki-

[46] G. N. Gill and G. M. Walton, *Adv. Cyclic Nucleotide Res.* **10**, 93 (1979).
[47] S. Mårdh, *Anal. Biochem.* **63**, 1 (1975).
[48] G. Taborsky, *Adv. Protein Chem.* **28**, 1 (1974).

nase,[32] and casein kinase.[49] In experiments with cyclic AMP-dependent protein kinase, the presence of bovine serum albumin keeps the activity of the usually highly diluted protein kinase reasonably constant during the incubation.[32] This is particularly important in kinetic studies. Another difficulty with kinetic experiments arises during measurement of the rate of phosphorylation of peptides with low K_m values. In order to measure the rate at sub-K_m concentrations of the peptide, the radiochemical purity of [^{32}P]ATP is particularly important. Even trace amounts of labeled impurities may give unacceptably high blank values.

Termination of the Reaction. The method for terminating the reaction depends on the further processing of the material. In kinetic experiments trichloroacetic acid[31] has been added to a final concentration of 10% or acetic acid[32] to a final concentration of 30%, and in the preparation of larger amounts of phosphopeptide, boiling for 3 min has been used.[50]

Analysis of Phosphorylation. Separation of [^{32}P]phosphopeptide from [^{32}P]ATP in kinetic experiments on purified protein kinase is achieved by electrophoresis[31] or by ion-exchange chromatography, e.g., on AG-1-X8, equilibrated with 30% acetic acid and packed in Pasteur pipettes.[32] In the former case, the trichloroacetic acid-containing sample is neutralized by five extractions with water-saturated ethyl ether. Electrophoresis is performed on Whatman 3 MM paper or on a thin layer of cellulose powder on a plastic base, e.g., Polygram Cel 300 (Machery-Nagel, Düren, Germany). Radioactive spots are located by autoradiography, cut out, and counted in empty scintillation vials (Čerenkov radiation). When isolated by an anion exchanger, the [^{32}P]phosphopeptide is eluted directly into an empty scintillation vial. The eluate is made alkaline with sodium hydroxide before counting, to prevent the evaporation of acetic acid in the scintillation counter.

Preparation of Phosphopeptides. For preparation purposes, the phosphopeptide is easily separated from unphosphorylated peptide and ATP by ion-exchange chromatography, preferably on CM-cellulose.[50] The phosphopeptide is easily located in the chromatogram by the addition of trace amounts of ^{32}P-labeled phosphopeptide to the sample loaded onto the column. More recently, HPLC systems with a capacity to separate a number of phosphopeptides from each other have been developed.[37]

Phosphorylation of Proteins *in Vivo* and in Intact Cells

Ideally, in an unprejudiced search for a physiologically relevant protein phosphorylation, intact animals injected with [^{32}P]orthophosphate

[49] B. E. Kemp, E. Benjamini, and E. G. Krebs, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1038 (1976).

[50] V. P. K. Titanji, Ö. Zetterqvist, and U. Ragnarsson, *FEBS Lett.* **78**, 86 (1977).

should be used. In practice, however, this approach presents a number of difficulties, one being the limits as to the specific radioactivity that can be used in such a system.

An alternative approach is to investigate isolated cells, obtained by collagenase treatment of an organ, for instance, the liver,[51] heart,[52] or fat pads.[53] If 4 MBq of [^{32}P]orthophosphate is added to a suspension of 3×10^6 hepatocytes in 1 ml of Krebs–Ringer solution, containing 1.18 μmol of orthophosphate, the specific radioactivity of the γ-phosphorus of [^{32}P]ATP in the cells will increase to a steady state of about 10 dpm/pmol in 30–40 min at 37°. Thus, the total amount of radioactivity handled may be considerably lower than that required in experiments on whole animals. Nevertheless, the specific radioactivity is sufficiently high to permit detection even of minor phosphoproteins.

There is a need for rapid isolation of [^{32}P]phosphoprotein to avoid losses of the [^{32}P]phosphate bound and for removal of other phosphoproteins in order to quantitate the ones under investigation. One method is to precipitate the protein from a homogenate of the ^{32}P-labeled cells by means of ammonium sulfate. The phosphorylation–dephosphorylation will then stop, owing to the lack of Mg^{2+} and ATP. A particular [^{32}P]phosphoprotein can also be isolated by immunoprecipitation, and the [^{32}P]phosphate incorporated can be measured directly in the immunoprecipitate. Another procedure is to lyse the cells in denaturing media that are used to prepare samples for one- or two-dimensional polyacrylamide gel electrophoresis.[8]

After the preincubation of the cells with radioactive phosphate, the excess can be removed by centrifugation and the incubation of the cells can continue in the presence or the absence of hormones that are known or presumed to affect the phosphorylation of the protein in question. In addition to samples for measurement of the incorporation of [^{32}P]phosphate, samples for determination of the specific radioactivity of the γ-phosphorus of the [^{32}P]ATP should be collected.[6]

Concluding Remarks

For a protein phosphorylation to be physiologically important, certain criteria have to be fulfilled.[54–56] A main criterion is that the phosphory-

[51] M. N. Berry and D. S. Friend, *J. Cell Biol.* **43**, 506 (1969).
[52] T. Powell and V. W. Twist, *Biochem. Biophys. Res. Commun.* **72**, 327 (1976).
[53] M. Rodbell, *J. Biol. Chem.* **239**, 375 (1964).
[54] E. G. Krebs, *in* "Endocrinology, Proceedings of the Fourth International Congress of Endocrinology," p. 17. Excerpta Medica, Amsterdam, 1973.
[55] E. G. Krebs and J. A. Beavo, *Annu. Rev. Biochem.* **48**, 923 (1979).
[56] H. G. Nimmo and P. Cohen, *Adv. Cyclic Nucleotide Res.* **8**, 145 (1977).

lation occurs *in vivo*. In order to study the *in vivo* phosphorylation of minor components, extremely high specific radioactivity of the endogenous ATP is required. This may be impossible to achieve with whole animals, in which case experiments on isolated cells will have to be accepted.

One way to establish that the phosphorylation detected in an *in vitro* system is the same as that observed *in vivo* or in intact cells is to determine the amino acid sequence near the phosphorylated amino acid residue. Alternatively, the pattern of [^{32}P]phosphopeptides obtained after enzymatic digestion of partial acid hydrolysis[17] may be investigated.

For many studies on protein phosphorylation, however, broken-cell preparations are employed. If a protein in a crude extract is phosphorylated by a purified protein kinase, it may be reasonable to assume that the phosphorylation is of physiological significance. However, artifactual phosphorylation may be obtained upon denaturation. It has been shown that a short segment of the peptide chain around the phosphorylatable amino acid contains the information that is sufficient for a significant rate of phosphorylation.[31] This points to the risk that partial denaturation of a protein will unmask a phosphorylatable site that is not exposed in the native state. It has been demonstrated, in fact, that denatured lysozyme, in contrast to the native enzyme, is phosphorylated by cyclic AMP-dependent protein kinase.[28]

Synthetic peptides, representing the phosphorylatable site of a protein, have proved to be invaluable tools in the elucidation of the mechanism of, and structural requirements in, protein phosphorylation. However, in the experience of the authors, peptides that are phosphorylatable by cyclic AMP-dependent protein kinase are extremely sensitive to peptidases of crude extracts. This is true also for the corresponding phosphopeptides.[57] Since the cleavage may separate the arginine residues from the remainder of the phosphopeptide, the initially basic phosphopeptide turns into an acid compound that may not be detected in an assay procedure designed to detect basic phosphopeptides. The use of peptide substrates in the determination of cyclic AMP-dependent protein kinase activity in crude extracts may therefore be subject to errors.

The number of enzymes and other identified proteins that are phosphorylated by protein kinases has been increasing rapidly in recent years. The table represents an attempt at a comprehensive list with respect to enzymes. With a few exceptions, only mammalian protein kinases and substrates are listed. Phosphorylations in which the substrate has been less well defined are mostly excluded. Despite the length of the list, a number of intracellular substrates of protein kinases are still unidentified.

[57] Ö. Zetterqvist and R.-M. Bålöw, *Biochem. Soc. Trans.* 9, 233 P (1981).

[6]

PROTEIN SUBSTRATES OF PROTEIN KINASES[a]

Protein kinase	Inhibitors and activators of the protein kinase	Protein substrate	Effect of the phosphorylation on the biological activity of the substrate	Amino acid sequence at the phosphorylated site	References[b]
cAMP-dependent protein kinase = glycogen synthase kinase I	cAMP (+)	Acetyl-CoA carboxylase	–		1
	Walsh inhibitor (–)	Actin			2
		ATP citrate-lyase		ThrAlaSerPPheSerGlu	3
		Cholesterol esterase	+		4, 5
		Cyclic nucleotide phosphodiesterase			6
		Fibrinogen		ThrThrArgArgSerPCysSerLys	7
		Filamin			8
		Fructose-1,6-bis-phosphatase	+	SerArgTyrSerPLeuProLeu	9
		Fructose-2,6-bis-phosphatase/fructose-6-P,2-kinase	+		10
			–		11
		Fructose-6-P kinase	–	SerArgLysArgSerPGlyGluAla	12, 13
		Glycogen synthase	–	ProArgArgAlaSerPCysThrSer	14
				LysArgSerAsnSerPValAspThr	14
				SerArgThrLeuSerPValSerSer	14
		G-substrate		ArgArgLysAspThrPProAlaLeu	15
		Guanylate cyclase	+		16
		Histone		ArgLysAlaSerPGlyProPro (H1)	17, 18
				ThrArgSerSerPArg (H2A)	17, 18
				LysArgSerPArgLysGluSerPTyr (H2B)	17, 18

(*continued*)

PROTEIN SUBSTRATES OF PROTEIN KINASES (*continued*)

Protein kinase	Inhibitors and activators of the protein kinase	Protein substrate	Effect of the phosphorylation on the biological activity of the substrate	Amino acid sequence at the phosphorylated site	References[b]
		HMG 14		LysArgLysValSerPSerAlaGlu	19
		Hormone-sensitive lipase/diglyceride lipase	+		20, 21
		Lipomodulin	–		22
		Myelin basic protein		GlyArgGlyLeuSerPLeuSerArg	23
				ArgHisArgAspThrPGlyIleLeu	23
				GlnArgHisGlySerPLysTyrLeu	23
		Myosin light-chain kinase	–		24, 25
		Na^+,K^+-ATPase			26
		Phenylalanine hydroxylase	+	SerArgLysLeuSerPAsxPheGly	27
		Phosphatase inhibitor I	+	ArgArgArgProThrPProAlaThr	14
		Phosphorylase kinase	(+)	PheArgArgLeuSerPIleSerThr (α-chain)	28
			+	LysArgSerGlySerP$^{Val}_{Ile}$TyrGlu (β-chain)	14
		Protamine		ArgArgSerSerSerPArgProIle	29
		Pyruvate kinase	–	LeuArgArgAlaSerP$^{Val}_{Leu}$AlaGlx	30
		RNA polymerase	+		31
		Self: Catalytic subunit		GluIleArgValSerPIleAsnGlu	32, 33
		Regulatory subunit (R_{II})	–	AspArgArgValSerPValCysAla	34

		RNA polymerase	+		31
		Self: Catalytic subunit		GluIleArgValSerPIleAsnGlu	32, 33
		Regulatory subunit (R_{II})	–	AspArgArgValSerPValCysAla	34
		Troponin I	+	ArgValArgMetSerPAlaAspAla	35
			–	ArgArgSerPAspArgAla	35, 36
		Tyrosine hydroxylase	+		37
cGMP-dependent protein kinase	cGMP (+)	ATP-citrate lyase			38
	Protein modulator (+)	cAMP-dependent protein kinase (regulatory subunit R_I)	–	ArgGlyAlaIleSerPAlaGluVal	34
		Fructose-1,6-bisphosphatase			39
		Glycogen synthase	–		39
		G-substrate		ArgArgLysAspThrPProAlaLeu	15
		Histone		ArgLysAlaSerPGlyProPro (H1)	18, 40
				LysArgSerPArgLysGluSerPTyr (H2B)	18, 40
		HMG 14		LysArgLysValSerPSerAlaGlu	19
		Hormone-sensitive lipase	+		41
		Phosphorylase kinase	+		42
		Pyruvate kinase	–		39
		Self	+		43
		Troponin I			44, 45
Casein kinase I = casein kinase S		Casein		ValAsnGluLeuSerPLysAspIle (αs_1)	46
				SerPGluGluAsnSerPLysLysThr (αs_2)	46
				SerPSerPGluGluSerPIleIleSerP (αs_2)	46
				SerPSerPGluGluSerPIleThrArg (β)	46

(continued)

PROTEIN SUBSTRATES OF PROTEIN KINASES (*continued*)

Protein kinase	Inhibitors and activators of the protein kinase	Protein substrate	Effect of the phosphorylation on the biological activity of the substrate	Amino acid sequence at the phosphorylated site	References[b]
		Glycogen synthase	–		46
		Phosphorylase kinase	+		28
		Phosvitin			47
		RNA-polymerase			48
		Self			46
		Spectrin			46
Casein kinase II = casein kinase TS = troponin T kinase = glycogen synthase kinase V = casein kinase G = casein kinase NII	Heparin (–)	Acetyl CoA carboxylase			49
	2,3-Diphosphoglycerate (–)	Calsequestrin			50
	Polyamines (+)	cAMP-dependent protein kinase (regulatory subunit R_{II})		AlaAspSerPGluSerPGluAspGlu	51
		Casein		GlySerPGluSerPThrPGluAspGln (αs_1)	46
				GluGlnLeuSerPThrPSerPGluGlu (αs_2)	46
				GluGlnGlnGlnThrPGluAspGlu (β)	46
					46
		eIF-2β		SerProHisGlnSerPGluAspGlu	47, 52
		Glycogen synthase	–		50
		Phosphatase inhibitor 1			47
		Phosvitin			53
		RNA polymerase	+		46
		Self			54
		Troponin T		AcSerPAspGluGluValGlu	
		Casein			46, 55

		Phosphatase inhibitor I			50
		Phosvitin			47
		RNA polymerase	+		53
		Self			46
		Troponin T		AcSerPAspGluGluValGlu	54
Mammary gland casein kinase		Casein			46, 55
Calcium- and phospholipid-activated protein kinase	Ca^{2+} (+) μM (−) mM	Fibrinogen			56
	Phospholipids (+)	Histone			57, 58
	Diolein (+)	Myelin basic protein			59
	Phorbol esters (+)	Protamine			58
	Heparin (−)	Self			58
	Melittin (−)	Troponin I			59
	Phenothiazines (−)	Troponin T			59
	Polyamines (−)	Calcium-transport ATPase			60
Phosphorylase kinase = glycogen synthase kinase II	Ca^{2+} (+)	Fructose-6-P,2 kinase	−		11
	Calmodulin (+)	Glycogen synthase	−	SerArgThrLeuSerPValSerSer	61
	Phenothiazines (−)	Phosphorylase	+	ArgLysGlnIleSerPValArgGly	62
		Self	+		63
		Troponin I		AlaIleThrPAlaArgArg	35
				AlaLeuSerPThrArgCys	35
		Troponin T		SerPAsnGluGluValGlu	35
				AlaLeu(SerP,Ser)MetGlyAla	35
				AsnTyr(SerP,Ser)Tyr	35
Glycogen synthase kinase III		cAMP-dependent protein kinase (regulatory subunit R_{II})		AlaArgSerPArgAlaSerPThrPro	51
		Glycogen synthase	−	ProArgProAlaSerPValPro- -ProSerPProSerLeuSerPArg	51
		Self			64
Glycogen synthase kinase IV		Glycogen synthase	−	SerArgThrLeuSerPValSerSer	49

(continued)

PROTEIN SUBSTRATES OF PROTEIN KINASES (*continued*)

Protein kinase	Inhibitors and activators of the protein kinase	Protein substrate	Effect of the phosphorylation on the biological activity of the substrate	Amino acid sequence at the phosphorylated site	References[b]
Glycogen synthase kinase	Ca^{2+} (+) Calmodulin (+) Phenothiazines (−)	Glycogen synthase Myosin light chain Self	−	SerArgThrLeuSerPValSerSer	65 66 66
Myosin light-chain kinase	Ca^{2+} (+) Calmodulin (+) Phenothiazines (−)	Myosin light chain	+	ArgAlaThrSerPAsnValPhe GlyGlySerSerPAsnValPhe	67 67
Hemin-dependent eIF-2α-kinase	Hemin (−)	eIF-2α	−		68
	N-Ethylmaleimide (+)	Self			69
Double stranded RNA-dependent protein kinase	Double-stranded RNA (+) [high concentration (−)]	eIF-2α Histone	−		70 69
	N-Ethylmaleimide (−)	Self			70
Pyruvate dehydrogenase kinase	Acetyl-CoA (+) NADH (+) Pyruvate (−) ADP (−)	Pyruvate dehydrogenase	−	TyrHisGlyHisSerPMetSerAsn- -ProGlyValSerPTyrArg GlyMetGlyThrSerPValGluArg	71 71
Acetyl-CoA carboxylase kinase		Acetyl-CoA carboxylase	+		72
Acetyl-CoA carboxylase kinase		Acetyl-CoA carboxylase	−		73

boxylase kinase		boxylase		
Acetyl-CoA carboxylase kinase		Acetyl-CoA carboxylase	–	73
		Histone		73
		Protamine		73
Not identified		Aminoacyl-tRNA synthetase	–	74
ATP citrate-lyase kinase		ATP citrate-lyase		75
Branched chain α-ketoacid dehydrogenase kinase		Branched chain α-ketoacid dehydrogenase	–	76, 77
Calcium and calmodulin-stimulated protein kinase	Ca^{2+} (+)	Tubulin		78
	Calmodulin (+)	Tryptophan hydroxylase	+	79
	Phenothiazines (–)	Tyrosine hydroxylase	+	79
Not identified		Glycerophosphate acyltransferase	–	80
Hydroxymethylglutaryl-CoA reductase kinase		Hydroxymethylglutaryl-CoA reductase	–	81
		Self		81
Hydroxymethylglutaryl-CoA reductase kinase kinase		Hydroxymethylglutaryl-CoA reductase kinase	+	81
Polyamine-dependent protein kinase	Polyamines (+)	Ornithine decarboxylase	–	82

(continued)

PROTEIN SUBSTRATES OF PROTEIN KINASES (*continued*)

Protein kinase	Inhibitors and activators of the protein kinase	Protein substrate	Effect of the phosphorylation on the biological activity of the substrate	Amino acid sequence at the phosphorylated site	References[b]
Rhodopsin kinase		Rhodopsin			83
		Self			83
Not identified		Tyrosine amino-transferase			84

[a] The protein kinases mentioned toward the end of the table have as yet been less extensively investigated. Autophosphorylation ("self") is included even if it has not been proved that the kinase preparation was free of other protein kinases. When known, the effect of the phosphorylation on the biological activity of the substrate is indicated (+, increase; −, decrease). The amino acid sequence at the phosphorylated site is given only if the sequence of an isolated phosphopeptide has been determined or if the amino acid composition of the phosphopeptide and amino acid sequence of the protein was known. Some minor phosphorylation sites have been omitted. Owing to space limitations, not all the reported phosphorylation sites in nonenzyme substrates, such as casein and protamine, are listed.

The references consist mainly of recent and comprehensive articles containing either original or reviewed data.

[b] Key to references: (*1*) R. W. Brownsey and D. G. Hardie, *FEBS Lett.* **120,** 67 (1980); (*2*) M. P. Walsh, S. Hinkins, and D. J. Hartshorne, *Biochem. Biophys. Res. Commun.* **102,** 149 (1981); (*3*) M. W. Pierce, J. L. Palmer, H. T. Keutmann, T. A. Hall, and J. Avruch, *J. Biol. Chem.* **257,** 10681 (1982); (*4*) G. S. Boyd and A. M. S. Gorban, *Mol. Aspects Cell. Regul.* **1,** 95 (1980); (*5*) J. C. Khoo, E. M. Mahoney, and D. Steinberg, *J. Biol. Chem.* **256,** 12659 (1981); (*6*) R. K. Sharma, T. H. Wang, E. Wirch, and J. H. Wang, *J. Biol. Chem.* **255,** 5916 (1980); (*7*) L. Engström, B. Edlund, U. Ragnarsson, U. Dahlqvist-Edberg, and E. Humble, *Biochem. Biophys. Res. Commun.* **96,** 1503 (1980); (*8*) D. Wallach, P. J. A. Davies, and I. Pastan, *J. Biol. Chem.* **253,** 4739 (1978); (*9*) P. Ekman and U. Dahlqvist-Edberg, *Biochim. Biophys. Acta* **662,** 265 (1981); (*10*) M. R. El-Maghrabi, E. Fox, J. Pilkis, and S. J. Pilkis, *Biochem. Biophys. Res. Commun.* **106,** 794 (1982); (*11*) C. S. Richards and K. Uyeda, *J. Biol. Chem.* **257,** 8854 (1982); (*12*) E. Furuya and K. Uyeda, *J. Biol. Chem.* **255,** 11656 (1980); (*13*) D.-H. Söling and I. A. Brand, *Curr. Top. Cell. Regul.* **20,** 107 (1981); (*14*) P. Cohen, *Nature* (*London*) **296,** 613 (1982); (*15*) A. Aitken, T. Bilham, P. Cohen, D. Aswad, and P. Greengard, *J. Biol. Chem.* **256,** 3501 (1981); (*16*) J. Zwiller, M.-O. Revel, and P. Basset, *Biochem. Biophys. Res. Commun.* **101,** 1381 (1981); (*17*) S. V. Shlyapnikov, A. A. Arutyunyan, S. N. Kurochkin, L. V. Memelova, M. V. Nesterova, L. P. Sashchenko, and E. S. Severin, *FEBS Lett.* **53,** 316 (1975); (*18*) D. B. Glass and E. G. Krebs, *Annu. Rev. Pharmacol. Toxicol.* **20,** 363 (1980); (*19*) G. M. Walton, J. Spiess, and G. N. Gill, *J. Biol. Chem.* **257,** 4661 (1982); (*20*) L. Berglund, J. C. Khoo, D. Jensen, and D. Steinberg, *J. Biol. Chem.* **255,** 5420 (1980); (*21*) N. Ö. Nilsson, P. Strålfors, G. Fredrikson, and P. Belfrage, *FEBS Lett.* **111,** 125 (1980); (*22*) F. Hirata, *J. Biol. Chem.* **256,** 7730 (1981); (*23*) P. Daile, P. R. Carnegie, and J. D. Young, *Nature* (*London*) **257,** 416 (1975); (*24*) M. A. Conti and R. S. Adelstein, *J. Biol. Chem.* **256,** 3178 (1981); (*25*) A. M. Edelman and E. G. Krebs, *FEBS Lett.* **138,** 293 (1982); (*26*) S. Mårdh, *Curr. Top. Membr. Trans.* **19,** 999 (1983); (*27*) M. Wretborn, E. Humble, U. Ragnarsson, and L. Engström, *Biochem. Biophys. Res. Commun.* **93,** 403 (1980); (*28*) T. J. Singh, A. Akatsuka, and K.-P. Huang, *J. Biol. Chem.* **257,** 13370 (1982);

250, 5301 (1981); (*16*) J. Zwiller, M.-O. Revel, and P. Bassel, *Biochem. Biophys. Res. Commun.* **101**, 1381 (1981); (*17*) S. V. Shiyapnikov, A. A. Arutyunyan, S. N. Kurochkin, L. V. Memelova, M. V. Nesterova, L. P. Sashchenko, and E. S. Severin, *FEBS Lett.* **53**, 316 (1975); (*18*) D. B. Glass and E. G. Krebs, *Annu. Rev. Pharmacol. Toxicol.* **20**, 363 (1980); (*19*) G. M. Walton, J. Spiess, and G. N. Gill, *J. Biol. Chem.* **257**, 4661 (1982); (*20*) L. Berglund, J. C. Khoo, D. Jensen, and D. Steinberg, *J. Biol. Chem.* **255**, 5420 (1980); (*21*) N. Ö. Nilsson, P. Strålfors, G. Fredrikson, and P. Belfrage, *FEBS Lett.* **111**, 125 (1980); (*22*) F. Hirata, *J. Biol. Chem.* **256**, 7730 (1981); (*23*) P. Daile, P. R. Carnegie, and J. D. Young, *Nature* (*London*) **257**, 416 (1975); (*24*) M. A. Conti and R. S. Adelstein, *J. Biol. Chem.* **256**, 3178 (1981); (*25*) A. M. Edelman and E. G. Krebs, *FEBS Lett.* **138**, 293 (1982); (*26*) S. Märdh, *Curr. Top. Membr. Trans.* **19**, 999 (1983); (27) M. Wretborn, E. Humble, U. Ragnarsson, and L. Engström, *Biochem. Biophys. Res. Commun.* **93**, 403 (1980); (*28*) T. J. Singh, A. Akatsuka, and K.-P. Huang, *J. Biol. Chem.* **257**, 13379 (1982); (*29*) S. Shenolikar and P. Cohen, *FEBS Lett.* **86**, 92 (1978); (*30*) L. Engström, *Mol. Aspects Cell. Regul.* **1**, 11 (1980); (*31*) E. G. Kranias, J. S. Schweppe, and R. A. Jungman, *J. Biol. Chem.* **252**, 6750 (1977); (*32*) Y. S. Chiu and M. Tao, *J. Biol. Chem.* **253**, 7145 (1978); (*33*) S. Shoji, D. C. Parmelee, R. D. Wade, S. Kumar, L. H. Ericsson, K. A. Walsh, H. Neurath, G. L. Long, J. G. Demaille, E. H. Fischer, and K. Titani, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 848 (1981); (*34*) R. L. Geahlen, D. F. Carmichael, E. Hashimoto, and E. G. Krebs, *Adv. Enzyme Regul.* **20**, 195 (1981); (*35*) S. V. Perry, *Biochem. Soc. Trans.* **7**, 593 (1979); (*36*) K. Yamamoto and I. Ohtsuki, *J. Biochem.* (*Tokyo*) **91**, 1669 (1982); (*37*) A. M. Edelman, J. D. Raese, M. A. Lazar, and J. D. Barchas, *J. Pharmacol. Exp. Ther.* **216**, 647 (1981); (*38*) G. D. Swergold, O. M. Rosen, and C. S. Rubin, *J. Biol. Chem.* **257**, 4207 (1982); (*39*) T. M. Lincoln and J. D. Corbin, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3239 (1977); (*40*) E. Hashimoto, M. Takeda, Y. Nishizuka, K. Hamana, and K. Iwai, *J. Biol. Chem.* **251**, 6287 (1976); (*41*) J. C. Khoo, P. J. Sperry, G. N. Gill, and D. Steinberg, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4843 (1977); (*42*) P. Cohen, *FEBS Lett.* **119**, 301 (1980); (*43*) J. L. Foster, J. Guttman, and O. M. Rosen, *J. Biol. Chem.* **256**, 5029 (1981); (*44*) D. K. Blumenthal, J. T. Stull, and G. N. Gill, *J. Biol. Chem.* **253**, 334 (1978); (*45*) T. M. Lincoln and J. D. Corbin. *J. Biol. Chem.* **253**, 337 (1978); (*46*) G. M. Hathaway and J. A. Traugh, *Curr. Top. Cell. Regul.* **21**, 101 (1982); (*47*) E. Itarte, M. A. Mor, A. Salavert, J. M. Pena, J. F. Bertomeu, and J. J. Guinovart, *Biochim. Biophys. Acta* **658**, 334 (1981); (*48*) M. E. Dahmus, *J. Biol. Chem.* **256**, 3332 (1981); (*49*) P. Cohen, D. Yellowless, A. Aitken, A. Donella-Deana, B. A. Hemmings, and P. J. Parker, *Eur. J. Biochem.* **124**, 21 (1982); (*50*) F. Meggio, A. Donella-Deana, and L. A. Pinna, *J. Biol. Chem.* **256**, 11958 (1981); (*51*) B. A. Hemmings, A. Aitken, P. Cohen, M. Rymond, and F. Hofmann, *Eur. J. Biochem.* **127**, 473 (1982); (*52*) C. Picton, J. Woodgett, B. A. Hemmings, and P. Cohen, *FEBS Lett.* **150**, 191 (1982); (*53*) D. A. Stetler and K. M. Rose, *Biochemistry* **21**, 3721 (1982); (*54*) L. A. Pinna, F. Meggio, and M. M. Dediukina, *Biochem. Biophys. Res. Commun.* **100**, 449 (1981); (*55*) J. C. Pascall, A. P. Boulton, and R. K. Craig, *Eur. J. Biochem.* **119**, 91 (1981); (*56*) P. Papanikolaou, E. Humble, and L. Engström, *FEBS Lett.* **143**, 199 (1982); (*57*) M. Castagna, Y. Takai, K. Kaibuchi, K. Sano, U. Kikkawa, and Y. Nishizuka, *J. Biol. Chem.* **257**, 7847 (1982); (*58*) U. Kikkawa, Y. Takai, R. Minakuchi, S. Inohara, and Y. Nishizuka, *J. Biol. Chem.* **257**, 13341 (1982); (*59*) N. Katoh, B. C. Wise, and J. F. Kuo, *Biochem. J.* **209**, 189 (1983); (*60*) M. Varsanyi and L. M. G. Heilmeyer, Jr., *FEBS Lett.* **131**, 223 (1981); (*61*) N. Embi, D. B. Rylatt, and P. Cohen, *Eur. J. Biochem.* **100**, 339 (1979); (*62*) C. Nolan, W. B. Novoa, E. G. Krebs, and E. H. Fischer, *Biochemistry* **3**, 542 (1964); (*63*) G. M. Carlson, P. J. Bechtel, and D. J. Graves, *Adv. Enzymol. Relat. Areas Mol. Biol.* **50**, 41 (1979); (*64*) B. A. Hemmings, D. Yellowlees, J. C. Kernohan, and P. Cohen, *Eur. J. Biochem.* **119**, 443 (1981); (*65*) J. R. Woodgett, N. K. Tonks, and P. Cohen, *FEBS Lett.* **148**, 5 (1982); (*66*) Z. Ahmad, A. A. DePaoli-Roach, and P. J. Roach, *J. Biol. Chem.* **257**, 8348 (1982); (*67*) Symposium Oct. 1981: Coordination of Metabolism and Contractibility by phosphorylation in cardiac, skeletal and smooth muscle, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **42**, 7 (1983); (*68*) R. Jagus, D. Chrouch, A. Konieczny, and B. Safer, *Curr. Top. Cell. Regul.* **21**, 35 (1982); (*69*) H. Grosfeld and S. Ochoa, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6526 (1980); (*70*) R. Petryshyn, D. H. Levin, and I. M. London, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6512 (1982); (*71*) L. J. Reed, *Curr. Top. Cell. Regul.* **18**, 95 (1981); (*72*) R. W. Brownsey, G. J. Belsham, and R. M. Denton, *FEBS Lett.* **124**, 145 (1981); (*73*) B. Lent and K.-H. Kim, *J. Biol. Chem.* **257**, 1897 (1982); (*74*) Z. Damuni, F. B. Caudwell, and P. Cohen, *Eur. J. Biochem.* **129**, 57 (1982); (*75*) S. Ramakrishna, D. L. Pucci, and W. B. Benjamin, *J. Biol. Chem.* **256**, 10213 (1981); (*76*) K. S. Lau, H. R. Fatania, and P. J. Randle, *FEBS Lett.* **144**, 57 (1982); (*77*) R. Paxton and R. A. Harris, *J. Biol. Chem.* **257**, 14433 (1982); (*78*) B. E. Burke and R. J. Delorenzo, *Brain Res.* **236**, 393 (1982); (*79*) T. Yamauchi, H. Nakata, and H. Fujisawa, *J. Biol. Chem.* **256**, 5404 (1981); (*80*) H. G. Nimmo, *Mol. Aspects Cell. Regul.* **1**, 135 (1980); (*81*) Z. H. Beg and H. B. Brewer, Jr., *Curr. Top. Cell. Regul.* **20**, 139 (1981); (*82*) G. D. Kuehn and V. J. Atmar, *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **41**, 3078 (1982); (*83*) R. H. Lee, B. M. Brown, and R. N. Lolley, *Biochemistry* **21**, 3303 (1982); (*84*) R. H. Stellwagen and K. K. Kohli, *Biochem. Biophys. Res. Commun.* **102**, 1209 (1981).

One reason for this is the fact that unequivocal identification of a substrate requires its extensive purification. When the substrate is an enzyme, such a purification generally calls for methods of stabilizing the enzyme. Since most phosphorylatable enzymes are labile, the identification of a substrate of a protein kinase may be a fairly difficult problem. In addition, the identification may be hampered by the fact that several particle-bound protein kinase activities are not very well characterized. It is even possible that there are protein kinases that still remain to be discovered.

[7] Assays for Regulatory Properties of Polyamine-Dependent Protein Kinase

By VAITHILINGAM SEKAR, VALERIE J. ATMAR, and GLENN D. KUEHN

[EC 2.7.1.37, ATP: protein phosphotransferase]

$$ATP^{4-} + \text{ornithine decarboxylase enzyme} \rightarrow ADP^{3-} + H^{+} + \text{phosphoornithine decarboxylase}^{2-}$$

The naturally occurring polyamines putrescine, spermidine, and spermine are synthesized in appreciable amounts in all living cells.

Putrescine: $H_2N\text{-}CH_2CH_2CH_2CH_2\text{-}NH_2$
Spermidine: $H_2N\text{-}CH_2CH_2CH_2CH_2\text{-}NH\text{-}CH_2CH_2CH_2\text{-}NH_2$
Spermine: $H_2N\text{-}CH_2CH_2CH_2\text{-}NH\text{-}CH_2CH_2CH_2CH_2\text{-}NH\text{-}CH_2CH_2CH_2\text{-}NH_2$

The combined charged cationic and aliphatic properties of the polyamines, under physiological conditions, suggest a wide variety of interactions that may occur between them and various cellular components. They can link through ionic forces to anionic nucleic acids, proteins, and phospholipids. The aliphatic character of their methylene group clusters indicates possible interactions with hydrophobic environments such as those occurring in membranes. Their amino-linked protons can hydrogen-bond to electronegative atoms. Thus, the polyamines have been implicated to act in a wide variety of cellular processes. Their participation has been invoked in virtually every phase of macromolecular biosynthesis.

The polyamines are synthesized in eukaryotes by a pathway that begins with the decarboxylation of the amino acid ornithine. The enzyme ornithine decarboxylase [L-ornithine carboxy-lyase, EC 4.1.1.17] catalyzes putrescine formation through pyridoxal phosphate-dependent de-

METHODS IN ENZYMOLOGY, VOL. 107

ISBN 0-12-182007-6

EXHIBIT 9

THE USE OF SYNTHETIC PEPTIDES FOR DEFINING THE SPECIFICITY OF TYROSINE PROTEIN KINASES

JOHN E. CASNELLIE and EDWIN G. KREBS

Howard Hughes Medical Institute Laboratories, University of Washington, School of Medicine, Seattle, Washington 98195

INTRODUCTION

The tyrosine protein kinases (ATP: protein phosphotransferase EC 2.7.1.37) encompass a large family of at least 10 distinct enzymes. The members of this family are related not only in their specificity for phosphorylating proteins on tyrosine residues but also in their biological function. Thus, all of these enzymes may have roles in regulating cell growth and differentiation. Moreover, the enzymes associated with RNA tumor viruses (at least in all cases thus far examined) show sequence homology, indicating that this group of tyrosine protein kinases may have evolved from a common ancestral protein. In order to understand more fully the biological roles of these enzymes it is necessary to understand more about their biochemical properties. One important property of any protein kinase is its substrate specificity. One would like to define those structural elements of the substrate, besides the phosphorylatable amino acid residue, that are important in the recognition of the substrate by the protein kinase. In the case of the tyrosine protein kinases there are two especially important reasons for wanting to study their specificities. One reason is that if the specificities of these enzymes are similar then it is possible that some or all of these enzymes are regulating the same biochemical pathways through their capacity to covalently modify the same regulatory proteins. A second reason is the desire to determine if the viral tyrosine protein kinases have a different specificity from their normal cellular counterparts which could explain the transforming properties of the viral enzymes.

Synthetic peptides provide a useful means for studying the specificity of protein kinases. They can be prepared in sufficiently large quantities to allow for the accurate determination of kinetic parameters, and their sequence can be varied at will, allowing one to probe the contribution of any residue to the recognition of the peptide by the protein kinase. The demonstration that synthetic peptides can serve as substrates for tyrosine protein kinases (1) provides the opportunity to use synthetic peptides to study the specificity of these enzymes. In this article we will discuss our work and that of other laboratories in this area. We shall preface this discussion by reviewing the

results of specificity studies with synthetic peptides of the cyclic nucleotide-dependent protein kinases. The specificity of these enzymes has been rather thoroughly explored in studies that have utilized on the order of 100 different synthetic peptides. It is thus timely to consider what we have learned from the work on the cyclic nucleotide dependent enzymes in terms of what we can expect to learn about the specificity of the tyrosine protein kinases from the use of synthetic peptide substrates.

MATERIALS AND METHODS

The LSTRA cell line was maintained in culture as described (1). The cells were washed once with phosphate buffered saline and then homogenized in a Dounce homogenizer using a buffer consisting of 5 mM HEPES, pH 7.4, 1 mM $MgCl_2$ and 5 mM 2-mercaptoethanol. Sucrose was then added to a concentration of 0.25 M and the homogenate spun at low speed to remove nuclei. The high speed pellet was then obtained and resuspended in 25 mM HEPES, pH 7.4, 5 mM 2-mercaptoethanol. This fraction was then extracted with Triton X-100 and used as a source of LSTRA tyrosine protein kinase.

Solid-phase peptide synthesis was carried out with the Beckman 9908 automated instrument as described (1). The peptides were purified on SP-Sephadex using a gradient of ammonium acetate and desalted by chromatography in 30% acetic acid on Sephadex G-10. Their composition was confirmed by amino acid analysis. The peptide phosphorylation reactions were carried out under conditions where the reaction is linear (2) using 10 mM $MgCl_2$ and 100 μM ATP (2000–5000 cpm/pmol) in the reaction mixture.

RESULTS AND DISCUSSION

Specificity of Cyclic Nucleotide-Dependent Protein Kinases

Phosphorylation of synthetic peptides by cyclic AMP-dependent protein kinase. The initial work on the specificity of cyclic AMP-dependent protein kinase was done by varying the structure of a heptapeptide, Leu-Arg-Arg-Ala-Ser-Leu-Gly (3, 4), whose sequence was based on the site of phosphorylation in pyruvate kinase. Studies have also been done using peptides with sequences based on the sequences at the sites of phosphorylation in denatured lysozyme (5), phosphorylase (4), cardiac troponin I (6) and histone H2B (7). The results of these studies are summarized in Table 1. In general it appears that peptides with the structure of the site of phosphorylation in pyruvate kinase are the optimal peptide substrates for the cyclic AMP-dependent protein kinase in terms of both a low K_m and high V_{max}.

The critical feature of these peptides is the presence of 2 vicinal arginines on the N terminal side of the serine; substitution of either of these residues, even by another basic residue such as lysine, results in peptides with much greater K_m values and reduced V_{max} values. Studies where the number of residues between the arginine and the serine was systematically varied have demonstrated that the position of these 2 arginines is also important (8). They must be separated from the serine by a single residue. Any other spacing results in peptides with dramatically less favorable kinetic parameters (8). As summarized in Table 1, given a peptide with the optimally located pair of arginines, substitution of the other residues is generally either without effect or results in peptides that are poorer substrates. It is sometimes difficult to evaluate the effect of substitutions at the other positions simply because the changes in kinetic parameters are often less dramatic than substitutions at positions 2 and 3. In addition the effects are not always consistently observed and thus may depend on the overall sequence of the peptide. One interesting result is that while the 2 arginines must be separated from the serine by another residue, there is a great deal of tolerance in terms of the identity of this residue. A wide variety of residues has been substituted at this position, including large hydrophobic ones such as tryptophan (9), without significantly altering the kinetic properties of the peptides.

Out of the large number of peptides that have been studied there is only one that has been found to be a relatively good substrate for the cyclic AMP-dependent protein kinase that does not have a pair of arginine residues located at positions 2 and 3 (Table 1). This is a decapeptide with the sequence, Lys-Arg-Lys-Glu-Ile-Ser-Val-Ala-Gly-Leu (4). This peptide has the sequence of

TABLE 1. SPECIFICITY OF cAMP-DEPENDENT PROTEIN KINASE TOWARD SYNTHETIC PEPTIDES NH_2-1-2-3-4-Ser-6-7-COOH

Position	Positive determinant	Negative determinant	Presence of negative determinant at site of phosphorylation in an intact protein	References
1	—			
2	arg	non basic residue*	yes	(3–6, 10)
3	arg	non basic residue*	yes	(3–6, 10)
4	—	—	—	
5	ser†	—	—	
6	—	pro	yes	(10, 11)
7	—	arg, lys	yes	(4, 6, 7)

*Substitution of the arginine residue by a lysine residue results in a peptide that is a poorer substrate. However the substitution of a lysine is more favorable than the substitution of a neutral amino acid such as alanine (4).

†Peptides in which the serine is replaced by threonine are much poorer substrates although threonine residues are phosphorylated in intact proteins (4, 10, 12).

the site of phosphorylation in phosphorylase except that the underlined alanine was substituted for the arginine that occurs in the protein sequence. (This site in the intact protein is not phosphorylated by the cyclic AMP-dependent protein kinase). The substitution of this arginine by the alanine had an extremely dramatic effect on the peptide's kinetic parameters. As noted in Table 1, the presence of basic residues in this position relative to the serine is a negative determinant. In the case of the phosphorylase sequence, substitution of the arginine at this location by an alanine caused the K_m to drop from 3900 μM to 36 μM and the V_{max} to increase several fold. Only a limited amount of work was done on varying the structure of this peptide. It appears, however, that while the N-terminal lysine is not important, the Arg-Lys residues that follow in the sequence are both essential for the peptide to have a low K_m and a high V_{max}. It would be interesting to know whether the identity of the 2 residues between the Arg-Lys residues and the serine has any influence on the ability of the peptide to serve as a substrate.

Phosphorylation of synthetic peptides by cGMP-dependent protein kinase. There is another member of the family of protein kinases regulated by cyclic nucleotides besides the cyclic AMP-dependent enzyme; this is the protein kinase that is regulated by cyclic GMP. These two cyclic nucleotide-dependent enzymes are related in that they are probably also derived from a common ancestral protein (13, 14). Although the biological role of the cyclic GMP-dependent enzyme is not precisely defined, the processes regulated by this enzyme are probably distinct from those regulated by the cyclic AMP-dependent protein kinase.

The specificity of the cyclic GMP-dependent protein kinase has been studied with synthetic peptides with sequences based on the sites of phosphorylation in histone H2B (7). The results of this work have shown that the specificity of this enzyme toward synthetic peptides is similar to that of the cyclic AMP-dependent protein kinase. Thus, in order for a synthetic peptide to serve as a good substrate for the cyclic GMP-dependent protein kinase, the peptide must again have a pair of basic residues on the N-terminal side of the phosphorylatable serine. In the peptides examined thus far the pair of basic residues must also be separated from the serine by a single "spacer" residue (15). The presence of a basic residue at position 7 in the scheme of Table 1 is also a negative determinant for the cGMP-dependent protein kinase. The 2 cyclic nucleotide-dependent enzymes differ in their specificity toward synthetic peptides in that the cGMP-dependent enzyme seems to have a preference for an Arg-Lys sequence as the 2 basic residues (15). In addition, the presence of basic residues immediately adjacent to the serine on either the amino or carboxyl side is a positive determinant for the cGMP-dependent protein kinase. Basic residues in these positions with the histone H2B sequence are either a neutral or a somewhat negative determinant for the cAMP-dependent protein kinase.

The similar specificity of the 2 cyclic nucleotide dependent protein kinases toward synthetic peptides is also reflected in their specificity toward protein substrates. These 2 enzymes will both phosphorylate many of the same proteins albeit at different rates. For example, cardiac troponin I is phosphorylated by both enzymes. The K_m values for the 2 enzymes are nearly identical, but the cyclic AMP-dependent protein kinase has a V_{max} value that is 12-fold greater than that of the cGMP-dependent protein kinase (Table 2). Since in the same cell the levels of cyclic AMP and cyclic GMP are under the control of different hormones and neurotransmitters, it seems likely that these 2 cyclic nucleotides regulate different biological pathways. Nonetheless, the overlapping specificity of the 2 kinases suggests the possibility that in some instances they may phosphorylate the same proteins *in vivo*. Consequently, there may be cases where activation of either kinase produces the same effect in the cell.

Sequences at sites in proteins phosphorylated by cyclic AMP-dependent protein kinase. An examination of sequences at the site of phosphorylation in intact proteins reveals numerous examples in which the amino acid sequence at the phosphorylation site in the protein has residues in positions that would act as negative determinants if they were present in synthetic peptide substrates. A good example of this is with cardiac troponin I. Table 2 compares the abilities of the 2 cyclic nucleotide dependent protein kinases to phosphorylate this protein or a synthetic peptide having the sequence of the site of phosphorylation. The intact protein is an excellent substrate for the cyclic AMP-dependent protein kinase and a moderately good substrate for the cyclic GMP-dependent protein kinase. The synthetic peptide, on the other hand, is an extremely poor substrate for either enzyme. It is clear from the sequence of the site of phosphorylation why the synthetic peptide is such a

TABLE 2. KINETIC PARAMETERS FOR CYCLIC AMP AND CYCLIC GMP-DEPENDENT PROTEIN KINASES FOR THE PHOSPHORYLATION OF TROPONIN I AND A PEPTIDE WITH THE SEQUENCE OF THE SITE OF PHOSPHORYLATION

Substrate	Cyclic AMP-dependent protein kinase K_m (μM)	Cyclic AMP-dependent protein kinase V_{max} (μmole/min/mg)	Cyclic GMP-dependent protein kinase K_m (μM)	Cyclic GMP-dependent protein kinase V_{max} (μmole/min/mg)	Reference
Troponin I	21	11	16	0.9	(16)
Site of phosphorylation peptide, Ala-Val-Arg-Arg-Ser-Asp-Arg-Ala*	1500	0.012	400	0.008	(6, 7)

*A synthetic peptide which included more residues of the sequence on the N-terminus did not have improved kinetic parameters for the cyclic AMP-dependent protein kinase (6).

poor substrate. Thus, there is an arginine residue on the carboxyl side of the serine at a position where this is a negative determinant (Table 1). Moreover, the pair of arginines on the amino terminal side of the serine is not separated from the serine by another residue which is also unfavorable. Modifications of the sequence of this peptide so that it becomes more like the site of phosphorylation in pyruvate kinase resulted in greatly improved kinetic parameters.

There is another notable example in which the intact proteins are much better substrates than peptides with the sequences of the sites of phosphorylation. Several proteins, including glycogen synthase and the β-subunit of phosphorylase kinase, have sequences at their sites of phosphorylation in which the basic residues have the pattern, lys-arg-x-x-ser (10, 17). This observation prompted a study with synthetic peptides in which the investigators varied the number of residues between the Lys-Arg pair and the serine (8). It was found that in a synthetic peptide the optimal location for the lys-arg pair was in a position separated from the serine by only one residue. In contrast to the intact proteins, the peptide with the lys-arg separated from the serine by 2 residues was a very poor substrate. Another study, however, reported that the kinetic parameters of this latter peptide could be improved by adding a 3rd basic residue to give the pattern, arg-x-lys-arg-x-x-ser (18). This result is intriguing but does not completely explain the differences in the kinetic parameters between the intact proteins and the synthetic peptides. Thus, while the β-subunit of phosphorylase kinase has an arginine at this position (10) in its sequence, the site of phosphorylation in glycogen synthase lacks a basic residue at this position (17).

The above examples clearly imply that the kinases recognize more than simply the primary sequence around the phosphorylatable residue in the protein. These enzymes must also recognize a specific secondary structure in their substrates and the presence of this structure is just as important as the presence of certain residues in the primary sequence. The consequence of this secondary structural requirement is that cyclic nucleotide dependent protein kinases have a much narrower specificity toward synthetic peptides than toward proteins in terms of the primary sequence around the site of phosphorylation. Small peptides are extremely flexible and therefore do not have a single well defined secondary structure in solution. The site of phosphorylation in the intact protein would have a much more rigid structure in comparison to the synthetic peptide.

Peptides such as those with the sequence of the site of phosphorylation in cardiac troponin I apparently do not have the ability to readily form the proper secondary structure that is present in the intact protein. The inability to form the proper secondary structure probably accounts for the fact that some peptides are extremely poor substrates even though they have the optimally located pair or arginines. For example, substitution of a proline residue

immediately adjacent to the serine on the carboxyl side in the pyruvate kinase peptide results in a peptide that is inert to phosphorylation by the cyclic AMP-dependent protein kinase (11). The presence of the proline residue at this position presumably prevents the peptide from assuming the secondary structure that is recognized by the enzyme.

In discussing the specificities of protein kinase toward synthetic peptides and proteins one is essentially dealing with the problem of how the kinetics of the enzyme differs with the various substrates. It is not possible to completely understand the specificities of these enzymes until we understand how the structure of the substrate influences the individual rate constants of the reaction. The fact that cyclic nucleotide-dependent protein kinases recognize a specific secondary structure in their substrates means that there is an additional kinetic step in the reaction pathway when the enzyme acts on a synthetic peptide in comparison to when the enzyme phosphorylates a protein. This additional kinetic step involves either the induction of the specific secondary structure in the peptide upon binding to the enzyme or the selection of a specific, pre-existing conformation that is present at a low concentration relative to the total number of peptide molecules present (19). Indeed, because of this extra kinetic step, the on-rate of the pyruvate kinase heptapeptide for the cyclic AMP-dependent protein kinase is significantly slower than expected from considerations of rates of diffusion, although this does not prevent this peptide from being an excellent substrate (19). On the other hand, the troponin I peptide may be a poor substrate because this additional kinetic step is too slow for this peptide relative to the overall rate of the enzyme reaction.

Conclusions from the work with the cyclic nucleotide-dependent protein kinases. The studies with synthetic peptides together with the data on the phosphorylation of intact proteins demonstrate some of the problems in precisely defining the specificities of protein kinases. Nonetheless, the results with synthetic peptides do allow for some important conclusions. Perhaps the most satisfying conclusion concernes the unequivocal demonstration that the presence of basic residues near the N-terminus of the phosphorylatable residue is a crucial factor for substrate recognition by both cyclic mucleotide-dependent protein kinases. This result has predictive value in that one can be certain that any site that is readily phosphorylated by either of these enzymes will have basic residues at this position. The idea that basic residues are important for substrate recognition by the cyclic AMP-dependent protein kinase originally arose from considerations of the sequences around the various sites of phosphorylation in proteins. Synthetic peptides provided an extremely useful system for definitively testing this idea. It is also clear that recognition of secondary structure is an important factor in the substrate specificities of the cyclic nucleotide-dependent protein kinases. This result was foreshadowed for the cyclic AMP-dependent protein kinase from results with

protein substrates. Thus comparisons of the sequences around the site phosphorylated in proteins showed that with the exception of the invariable presence of basic residues on the N-terminal side of the phosphorylated residue (and even for these residues there is not a rigorous pattern as to their precise location relative to the phosphorylated residue) there is very little sequence homology among the various sites. This immediately suggests that primary structure is not the only factor in determining the sites phosphorylated by this enzyme. Finally, the studies with synthetic peptides also reveal the remarkable similarity in the specificity of the 2 cyclic nucleotide dependent protein kinases both in terms of primary and secondary structure recognition. As noted above, this result has potentially important implications concerning the biological functions of these 2 enzymes.

Specificity Studies on Tyrosine Protein Kinases

Studies using synthetic peptides. Initially, tyrosine protein kinase activity was detected as a result of the phosphorylation of antibody in immunoprecipitates and as a result of autophosphorylation reactions (20-23). The need to obtain specific substrates for these enzymes so that their activity could be measured in a conventional manner was an initial impetus for testing the ability of synthetic peptides to be phosphorylated by these enzymes (1). The synthetic peptides that were used in the initial studies had sequences that were based on the site of apparent autophosphorylation in $pp60^{src}$, the tyrosine protein kinase from Rous sarcoma virus. This was the first known sequence of a potentially physiologically relevant site of tyrosine phosphorylation. The sequences of this and other sites of autophosphorylation in different tyrosine protein kinases are shown in Table 3. The notable feature of these sequences is the presence of acidic residues on the amino side of the tyrosine. When the sequence at the site of phosphorylation in $pp60^{src}$ was determined, it was immediately suggested that these acidic

TABLE 3. SEQUENCES AT SITES OF AUTOPHOSPHORYLATION IN TYROSINE PROTEIN KINASES

Enzyme	Sequence	References
$pp60^{src}$, $pp90^{gag\text{-}yes}$, LSTRA	-Arg-Leu-Ile- Glu-Asp-Asn-Glu-Tyr*-Thr-Ala-Arg	(24, 25)
$pp130^{gag\text{-}fps}$	-Arg-Gln-Glu-Glu-Asp-Gly-Val-Tyr*-Ala-Ser-Thr	(26)
$pp87^{gag\text{-}fes}$	-Arg-Glu-Glu-Ala-Asp-Gly-Val-Tyr*-Ala-Ala-Ser	(27)
$pp120^{gag\text{-}abl}$	-Arg-Leu-Met-Thr-Gly-Asp-Thr-Tyr*-Thr-Ala-His	(28)

The acidic residues are underlined and the phosphorylated tyrosine is indicated with an asterisk(*).

residues might be important for the recognition of this site by the enzyme responsible for the phosphorylation (24), taking as a model the important role that basic residues have for the recognition of sites of phosphorylation of the cyclic nucleotide-dependent protein kinases. The work with synthetic peptides has provided some experimental verification of this concept.

The initial demonstration that a tyrosine protein kinase can phosphorylate a synthetic peptide was done using a tyrosine protein kinase that is present in the lymphoma cell line, LSTRA (1). This cell line has an elevated level of a tyrosine protein kinase with a molecular weight of 58,000, which appears to be distinct from the enzymes associated with the various RNA tumor viruses. It was initially shown that the enzyme from LSTRA cells would phosphorylate a synthetic peptide with the sequence of the site of tyrosine phosphorylation in $pp60^{src}$ (SRC-peptide) although the K_m for this peptide was rather high. Since this peptide is negatively charged the phosphorylated form of the peptide is difficult to separate from the γ-[^{32}P]ATP. This technical difficulty was overcome by synthesizing a peptide (RR-SRC-peptide) that contained the $pp60^{src}$ sequence but also had additional arginine residues at the N-terminus so that the phosphorylated peptide had a net positive charge in dilute acid. Under these conditions the phosphorylated peptide would then bind to phosphocellulose whereas γ-[^{32}P]ATP would not; this provided a simple and specific assay for tyrosine protein kinase activity (1).

There are 3 acidic residues on the N-terminal side of the tyrosine in the $pp60^{src}$ sequence. As demonstrated in Table 4, replacement of any one of these residues by an alanine residue did not have a significant effect on the ability of the LSTRA enzyme to phosphorylate the resultant peptide. However, replacement of all 3 acidic residues by neutral ones resulted in a large drop in the V_{max}. A peptide that had basic residues on the amino side of the tyrosine was a still poorer substrate. These results indicate that the presence of 1 or 2 acidic residues on the amino side of the tyrosine is a favorable determinant for the LSTRA protein kinase.

Results similar to those in Table 4 have also been found for the EGF-receptor tyrosine protein kinase (29, 30) using these and other peptides, although the K_m of this enzyme for the RR-SRC peptide was considerably lower than that found with the LSTRA enzyme. Again replacement of any 1 of the 3 acidic residues in the RR-SRC peptide did not have a profound effect on the kinetic parameters of the EGF-receptor protein kinase for the resultant peptides. However, the enzyme had little capacity to phosphorylate peptides with only basic amino acids on the amino side of the tyrosine. Thus acidic residues on the N-terminal side of the tyrosine appear to be important for a peptide to be an optimal substrate for the EGF-receptor protein kinase.

Another enzyme for which there is some indication that acidic residues are required for recognition of peptide substrates is the insulin receptor tyrosine protein kinase (31, 32). This enzyme will also phosphorylate the RR-SRC

TABLE 4. SPECIFICITY OF TYROSINE PROTEIN KINASE FROM LSTRA CELLS TOWARD SYNTHETIC PEPTIDES

Peptide*	K_m (mm)	V_{max} (nmoles/min/mg)
Arg-Arg-Leu-Ile-Glu-Asp-Ala-Glu-Tyr-Ala-Ala-Arg-Gly	2.0	8.0
Arg-Arg-Leu-Ile-Glu-Asp-Ala-Ala-Tyr-Ala-Ala-Arg-Gly	4.6	9.1
Arg-Arg-Leu-Ile-Glu-Ala-Ala-Glu-Tyr-Ala-Ala-Arg-Gly	2.7	5.5
Arg-Arg-Leu-Ile-Ala-Asp-Ala-Glu-Tyr-Ala-Ala-Arg-Gly	1.3	5.5
Glu-Asp-Ala-Glu-Tyr-Ala-Ala-Arg-Arg-Gly	3.4	9.0
Gly-Gly-Ala-Gly-Tyr-Ala-Ala-Arg-Arg-Gly	3.2	1.1
Leu-Arg-Arg-Ala-Tyr-Leu-Gly	6.7	0.6

The underlined residues indicate that a neutral amino acid has been substituted for an acidic residue that is present in the parent R-R-SRC peptide.

peptide (31, 32) although again the K_m is in the millimolar range. In addition the insulin receptor kinase will phosphorylate peptides with the sequence of the hormone, angiotensin (31). The physiological significance of this latter observation is dubious, but peptides with sequences based on the sequence of angiotensin contain tyrosine as the only phosphorylatable amino acid and are readily available, hence they are useful for specificity studies. In the limited study that was done it was found that the insulin receptor kinase will phosphorylate the angiotensin II peptide (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe) almost as well as it will phosphorylate the RR-SRC peptide. Angiotensin III, which is identical in sequence to angiotensin II, except that it is missing the N-terminal aspartic acid, was almost inert to phosphorylation by the insulin receptor kinase. Thus the presence of this 1 acidic residue was essential if the peptide was to be a substrate for this enzyme. This result is similar to that of a second study in which a peptide with only basic residues on the N-terminal side of the tyrosine was an extremely poor substrate in comparison to the RR-SRC peptide (32).

An investigation studying the capacity of the viral enzyme, $pp90^{gag\text{-}yes}$ (33) to phosphorylate the SRC peptide also found evidence that acidic residues were a factor in substrate recognition (33). This enzyme will also phosphorylate the SRC peptide but again with a K_m that is in the millimolar range. A deletion peptide that was missing the glutamic acid residue 4 residues from the amino end of the tyrosine was a much poorer substrate than the intact SRC peptide. The K_m of the $pp90^{gag\text{-}yes}$ for the deletion peptide was approximately 8-fold greater than for the parent SRC peptide, again demonstrating the potential importance of acidic residues for recognition of peptide substrates. The effect of deleting the other 2 acidic residues was not studied.

While the 4 enzymes discussed above all showed some indication of a preference for the presence of acidic residues on the amino side of the tyrosine in synthetic peptide substrates, the data are less clear for a fifth member of the

A CELLS

V_{max} les/min/mg)
8.0
9.1
5.5
5.5
9.0
1.1
0.6

for an acidic

In addition
equence of
f this latter
equence of
cid and are
the limited
kinase will
e-His-Pro-
ngiotensin
missing the
the insulin
ential if the
to that of a
N-terminal
to the RR-

ag-yes (33) to
dues were a
orylate the
A deletion
the amino
C peptide.
tely 8-fold
e potential
. The effect

ation of a
he tyrosine
mber of the

tyrosine protein kinase family, namely pp60^{src}. In a study using partially purified preparation of this kinase it was found that the enzyme had a rather broad specificity toward synthetic peptides. The pp60^{src} protein kinase could readily phosphorylate several angiotensin analogues that did not have acidic residues on the amino side of the tyrosine (34). In fact, several of these analogues that did not have acidic residues were better substrates than the peptide with the sequence of the site of phosphorylation in pp60^{src}. Thus the presence of acidic residues may not be a critical factor for substrate recognition by this protein kinase.

Although the above studies provided an indication that acidic residues on the N-terminal side of the tyrosine are a factor in substrate recognition by several of the tyrosine protein kinases, the changes in kinetic parameters seen with the various peptide substrates were often rather small. This is especially true when one compares how the presence or absence of basic residues affects the kinetic parameters of synthetic peptides for the cyclic AMP-dependent protein kinases. The presence of acidic residues *per se* may be of a more limited importance, since all the tyrosine protein kinases showed K_m values in the millimolar of tenth millimolar range for even the best peptide substrates.

Sequences at sites of tyrosine phosphorylation in proteins. There is very little data on the sequences phosphorylated in proteins by these enzymes, especially for physiological substrates. The only sequences known for sites of tyrosine phosphorylation in physiological substrates are the sites of auto-phosphorylation in some of the tyrosine protein kinases themselves. There have been a few studies on sites phosphorylated in proteins *in vitro*. The most detailed study done thus far has been on the sites phosphorylated in smooth muscle myosin light chain by the EGF-receptor protein kinase (35). This protein is a relatively good substrate for this enzyme. When the sites of phosphorylation were examined, it was found that there were 2 sites. Sequence determination showed that one site did indeed have a number of acidic residues on the N-terminal side of the tyrosine. However, the second site did not have any acidic or other charged residues in this position. The site that had acidic residues was phosphorylated at a faster rate but the difference was only a few fold. Thus the presence of acidic residues is not essential for a protein to be a substrate for the EGF-receptor protein kinase.

Another example in which the sequence of a site of phosphorylation in an *in vitro* substrate has been examined is the site phosphorylated in HLA-antigen by pp60^{src} (36). Unfortunately, no kinetic parameters were determined for this reaction; however, there was only one site of tyrosine phosphorylation and again this tyrosine did not have acidic residues near its N-terminal side.

Conclusions concerning the specificity of tyrosine protein kinases. The work carried out thus far demonstrates that while acidic residues are a factor in substrate recognition, especially for synthetic peptides, the presence of acidic residues in the primary sequence near the tyrosine is not an absolute

determinant for the recognition of sites of phosphorylation by the tyrosine protein kinases. It seems probable that secondary structure will also be important for the recognition of sites of phosphorylation by these ensymes. It may be that the presence of acidic residues is more important for the recognition of peptide substrates. Synthetic peptides will prove useful in comparing the specificity of the various tyrosine protein kinases and how their specificity is dependent on the primary structure of the substrate. Progress in this field will be greatly aided by the availability of purified enzymes and the identification of their physiological substrates.

SUMMARY

The tyrosine protein kinases are a large family of enzymes that may be involved in regulating cell growth and differentiation. An important property of these enzymes is their substrate specificity. Defining the specificity of these enzymes will contribute to a greater understanding of their biological functions. Synthetic peptides provide a useful means for studying the specificities of protein kinases. The utility of synthetic peptides for specificity studies is exemplified by the results obtained with the cyclic nucleotide-dependent protein kinases. A large number of synthetic peptides have been tested as substrates for these enzymes. There are three important conclusions from this work. First, in terms of primary sequence, the two cyclic nucleotide-dependent protein kinases both require the presence of a pair of basic residues on the N-terminal side of the phosphorylatable residue. Second, recognition of a specific secondary structure in the substrate is an equally important factor in the substrate specificities of these enzymes. Third, the specificities of the two cyclic nucleotide-dependent protein kinases are remarkably similar in terms of both primary and secondary structure recognition.

Sequences at the sites of tyrosine phosphorylation often show the presence of numerous acidic residues on the N-terminal side of the tyrosine. This result led to the suggestion that these acidic residues might be important for the recognition of these sites by the tyrosine protein kinases. Specificity studies using synthetic peptides have provided some experimental verification of this concept. In the case of four tyrosine protein kinases, LSTRA cell tyrosine protein kinase, epidermal growth factor receptor kinase, insulin receptor kinase and $pp60^{gag\text{-}yes}$ tyrosine protein kinase, the presence of acidic residues on the N-terminal side of the tyrosine in a synthetic peptide was a favorable determinant. In the case of a fifth member of the tyrosine protein kinase family, namely $pp60^{src}$ kinase, the data are less clear that the presence of acidic residues is involved in substrate recognition. This enzyme readily phosphorylated several peptides that did not have acidic residues on the N-terminal side of the tyrosine.

Although there is some indication that the presence of acidic residues on the

N-terminal side of the tyrosine is a factor in substrate recognition by several of the tyrosine protein kinases, the changes in kinetic parameters with the various peptide substrates were rather small. In addition, some sites phosphorylated in proteins do not have any acidic residues on the N-terminal side of the phosphorylated tyrosine residue. The presence of acidic residues on the primary sequence near the tyrosine is not an absolute determinant for the recognition of sites of phosphorylation by the tyrosine protein kinases. As in the case of the cyclic nucleotide-dependent protein kinases, secondary structure is probably also important for the recognition of sites of phosphorylation by the tyrosine protein kinases.

ACKNOWLEDGEMENTS

This work was supported by a National Institutes of Health National Research Service Award GM 07242-01 to J.E.C.

REFERENCES

1. J. E. CASNELLIE, M. L. HARRISON, L. J. PIKE, K. E. HELLSTROM and E. G. KREBS, Phosphorylation of synthetic peptides by a tyrosine protein kinase from the particulate fraction of a lymphoma cell line, *Proc. Natl. Acad. Sci. USA* 79, 282–286 (1982).
2. J. E. CASNELLIE, M. L. HARRISON, K. E. HELLSTROM and E. G. KREBS, A lymphoma cell line expressing elevated levels of tyrosine protein kinase activity, *J. Biol. Chem.* 258, 10738–10742 (1983).
3. O. ZETTERQUIST, U. RAGNARSSON, E. HUMBLE, L. BERGLUND and L. ENGSTROM, The minimum substrate of cyclic AMP-stimulated protein kinase, as studied by synthetic peptides representing the phosphorylation site of pyruvate kinase (type L) of rat liver, *Biochem. Biophys. Res. Comm.* 70, 696–703 (1976).
4. B. E. KEMP, D. J. GRAVES, E. BENJAMINI and E. G. KREBS, Role of multiple basic residues in determining the substrate specificity of cyclic AMP-dependent protein kinase, *J. Biol. Chem.* 252, 4888–4894 (1977).
5. B. E. KEMP, E. BENJAMINI and E. G. KREBS, Synthetic hexapeptide substrates and inhibitors of 3':5'-cyclic AMP-dependent protein kinase, *Proc. Natl. Acad. Sci. USA* 73, 1038–1042 (1976).
6. B. E. KEMP, Phosphorylation of synthetic peptide analogs of rabbit cardiac troponin inhibitory subunit by the cyclic AMP-dependent protein kinase, *J. Biol. Chem.*, 254, 2638–2642 (1979).
7. D. B. GLASS and E. G. KREBS, Phosphorylation by guanosine 3':5'-monophosphate-dependent protein kinase of synthetic peptide analogs of a site phosphorylated in histone H2B, *J. Biol. Chem.* 257, 1196–1200 (1982).
8. J. R. FERAMISCO, D. B. GLASS and E. G. KREBS, Optimal spatial requirements for the location of basic residues of peptide substrates for the cyclic AMP-dependent protein kinase, *J. Biol. Chem.* 235, 4240–4245 (1980).
9. D. E. WRIGHT, E. S. NOIMAN, P. B. CHOCK and V. CHAU, Fluorimetric assay for adenosine 3', 5'-cyclic monophosphate dependent protein kinase and phosphoprotein phosphatase activities, *Proc. Natl. Acad. Sci. USA* 378, 6048–6050 (1981).
10. S. SHENOLIKAR and P. COHEN, The substrate specificity of cyclic AMP-dependent protein kinase, *FEBS Lett.* 86, 92–98 (1978).
11. J. GRANOT, A. S. MILDVAN, H. N. BRANSON, N. THOMAS and E. T. KAISER, Nuclear magnetic resonance studies of the conformation and kinetics of the peptide-substrate at the active site of bovine heart protein kinase, *Biochemistry* 20, 602–610 (1981).

12. B. E. KEMP, I. D. RAE, E. MINASIAN and S. J. LEACH, Synthetic peptide substrates for protein kinases: structure and biological function, pp 169–172 in *Peptides, Structure and Biological Function* (E. GROSS and J. MEINHOFER, eds.) Pierce Chemical Co., Rockford, Ill. (1979).
13. T. M. LINCOLN and J. D. CORBIN, Adenosine 3':5'-cyclic monophosphate- and guanosine 3': 5' cyclic monophosphate-dependent protein kinases: possible homologous proteins, *Proc. Natl. Acad. Sci. USA* 74, 3239–3243 (1977).
14. E. HASHIMOTO, K. TAKIO and E. G. KREBS, Amino acid sequence at the ATP-binding site of cGMP dependent protein kinase, *J. Biol. Chem.* 257, 727–733 (1982).
15. D. B. GLASS, L. J. McFANN, M. D. MILLER and C. E. ZELIG, Interaction of cyclic-GMP-dependent protein kinase with phosphate-accepting proteins and peptides, pp 267–291 in *Cold Spring Harbor Conferences on Cell Proliferation.* 8 (O. M. ROSEN and E. G. KREBS, eds.) Cold Spring Harbor Laboratory (1981).
16. D. K. BLUMENTHAL, J. T. STULL and G. N. GILL, Phosphorylation of cardiac troponin by guanosine 3': 5'-monophosphate-dependent protein kinase, *J. Biol. Chem.* 253, 334–336 (1978).
17. T. S. HUANG, J. R. FERAMISCO, D. B. GLASS and E. G. KREBS, Specificity considerations relevant to protein kinase activation and function, pp. 449–461 in *Miami Winter Symposia,* 16 (T. R. RUSSELL, K. BREW, H. FARBER and J. SCHULTZ, eds.) Academic Press, New York, (1979).
18. O. ZETTERQUIST and U. RAGNARSSON, The structural requirements of substrates of cyclic AMP-dependent protein kinase, *FEBS Letters* 139, 287–290 (1982).
19. J. GRANOT, A. S. MILDVAN and E. T. KAISER, Studies of the mechanism of action and regulation of cAMP-dependent protein kinase, *Arch. Biochem. Biophys.* 205, 1–17, (1980).
20. M. S. COLLETT and R. L. ERIKSON, Protein kinase activity associated with the avian sarcoma virus src gene product, *Proc. Natl. Acad. Sci. USA* 75, 2021–2024 (1978).
21. O. N. WITTE, A. DASGUPTA and D. BALTIMORE, Abelson murine leukemia virus protein is phosphorylated *in vitro* to form phosphotyrosine, *Nature* 283, 826–831 (1980).
22. S. KAWAI, M. YOSHIDA, K. SEGAWA, H. SUGIYAMA, R. ISHIZAKI and K. TOYOSHIMA, Characterization of Y73, an avian sarcoma virus: A unique transforming gene and its product, a phosphopolyprotein with protein kinase activity, *Proc. Natl. Acad. Sci. USA* 77, 6199–6203 (1980).
23. M BARBACID, K. BEEMON and S. G. DEVARE, Origin and functional properties of the major gene product of the Synder-Theilen strain of feline sarcoma virus, *Proc. Natl. Acad. Sci. USA* 77, 5158–5162 (1980).
24. J. C. NEIL, J. GHYSDAEL, P. K. VOGT and J. E. SMART, Homologous tyrosine phosphorylation sites in transformation-specific gene products of distinct avian sarcoma viruses, *Nature* 291, 675–677 (1981).
25. J. E. CASNELLIE, M. L. HARRISON, K. E. HELLSTROM and E. G. KREBS, A lymphoma protein with an *in vitro* site of tyrosine phosphorylation homologous to that in $pp60^{src}$, *J. Biol. Chem.* 257, 13877–13879 (1982).
26. M. SHIBUYA and H. HANAFUSA, Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses, *Cell* 30, 787–795 (1982).
27. A. HAMPI, I. LAPREVOTTE, F. GALIBERT, L. A. FEDELE and C. J. SHERR, Nucleotide sequences of feline retroviral oncogenes (v-fes) provide evidence for a family of tyrosine-specific protein kinase genes. *Cell* 30, 775–785 (1982).
28. J. GOFFEN, N. HEISTERKAMP, F. H. REYNOLDS and J. R. STEPHENSON, Homology between phosphotyrosine acceptor site of human c-abl and viral oncogene products, *Nature* 304, 167–169 (1984).
29. L. J. PIKE, B. GALLIS, J. E. CASNELLIE, P. BORNSTEIN and E. G. KREBS, Epidermal growth factor stimulates the phosphorylation of synthetic tyrosine-containing peptides by A431 cell membranes, *Proc. Natl. Acad. Sci. USA* 79, 1443–1447 (1982).
30. C. ERNEUX, S. COHEN and D. L. GARBERS, The kinetics of tyrosine phosphorylation by the purified epidermal growth factor receptor kinase of A 431 cells. *J. Biol. Chem.* 258, 4137–4132 (1983).
31.
32.
33.
34.
35.
36.

substrates for
Structure and
o., Rockford,

osphate- and
homologous

ATP-binding

ion of cyclic-
peptides, pp
EN and E. G.

n of cardiac
l. Chem. 253,

S, Specificity
61 in *Miami*
IULTZ, eds.)

substrates of

of action and
1–17, (1980).
ith the avian
1978).
ukemia virus
5–831 (1980).
AKI and K.
transforming
c. Natl. Acad.

properties of
s, *Proc. Natl.*

ous tyrosine
ian sarcoma

KREBS, A
ous to that in

rcoma virus:
ther sarcoma

J. SHERR,
or a family of

PHENSON,
al oncogene

S, Epidermal
g peptides by

sphorylation
l. Chem. 258,

31. L. A. STADMAUER and O. M. ROSEN, Phosphorylation of exogenous substrates by the insulin receptor-tyrosine protein kinase, *J. Biol. Chem.* 258, 6682–6685 (1983).
32. M. KASUGA, Y. F. YAMAGUCHI, D. L. BLITHE, M. F. WHITE and C. R. KAHN, Characterization of the insulin receptor kinase purified from human placenta membranes, *J. Biol. Chem.* 258, 10973–10980 (1983).
33. T. HUNTER, Synthetic peptide substrates for a tyrosine protein kinase, *J. Biol. Chem.* 257, 4843–4848 (1982).
34. T. W. WONG and A. R. GOLDBERG, *In vitro* phosphorylation of angiotensin analogs by tyrosyl protein kinases, *J. Biol. Chem.* 258, 1022–1025 (1983).
35. B. GALLIS, A. M. EDELMAN, J. E. CASNELLIE and E. G. KREBS, Epidermal growth factor stimulates tyrosine phosphorylation of myosin regulatory light chain from smooth muscle, *J. Biol. Chem.*, 258, 13089–13093 (1983).
36. B. C. GUILD, R L. ERIKSON and J. L. STROMINGER, HLA-A2 and HLA-B7 antigens are phosphorylated *in vitro* by Rous sarcoma virus kinase ($pp60^{v\text{-}src}$) at a tyrosine residue encoded in a highly conserved exon of the intracellular domain, *Proc. Natl. Acad. Sci. USA* 80, 2894–2898 (1983).

EXHIBIT 10

13

Identifying protein kinase substrates by expression screening with solid-phase phosphorylation

RIKIRO FUKUNAGA and TONY HUNTER

1. Introduction

The biological activities of protein kinases are evoked through phosphorylation of their substrate proteins. Phosphorylation of target proteins causes changes in their structure, stability, enzymatic activity, ability to interact with other molecules, or subcellular localization, leading to regulation of a wide variety of cellular processes. Indeed, the identification of physiological targets has been a high priority ever since the first protein kinase was purified. However, the conventional approach of purifying substrate proteins by biochemical techniques is laborious and time consuming, and is especially difficult in the case of scarce proteins. Among the approaches developed to identify protein kinase substrates in a systematic manner are various techniques for determining consensus phosphorylation site sequences using oriented peptide libraries (1) (see Chapter 16), interaction screening for protein substrates by the far-Western method (2), or the yeast two-hybrid system (3, 4) (see Chapter 14).

In this chapter, we describe an alternative screening method for identifying protein kinase substrates. This technique, termed 'phosphorylation screening', utilizes a cDNA-expressing λ phage library, with phosphorylation of expressed proteins in the solid-phase. Application of this strategy to the ERK1 MAP kinase system resulted in the isolation of several cDNAs encoding both known and novel substrates (5). We have also used phosphorylation screening to identify substrates for cyclin E/Cdk2 (6), and we expect that it will be generally applicable for direct identification of physiological targets of various protein kinases. A similar phosphorylation screening method has recently been developed for the identification of protein–tyrosine kinase substrates using a λgt11 cDNA expression library (7).

2. Phosphorylation screening of a phage expression library

In this method, a λgt11-like phage expression library is screened using *in vitro*, solid-phase phosphorylation with [γ-^{32}P]ATP and the soluble protein kinase of interest. The method may be summarized as follows:

(a) A cDNA library is prepared using λGEX5 phage vector, in which a cDNA is inserted downstream of the glutathione-*S*-transferase (GST) coding region (5, 8).

(b) Phage plaques of the cDNA library are formed on agar plates, and GST-fused recombinant proteins expressed in the plaques are transferred and immobilized on nitrocellulose filters.

(c) The plaque filters are incubated with a purified, active protein kinase in the presence of [γ-^{32}P]ATP to allow solid-phase phosphorylation. The phosphorylated plaques are visualized by autoradiography.

(d) The positive clones are characterized by cDNA sequencing, and *in vitro* and *in vivo* analyses, to identify physiological substrates for the protein kinase.

The use of a solid-phase phosphorylation screening protocol is based on findings that cellular proteins immobilized on a membrane filter can be phosphorylated by a soluble protein kinase with similar specificity to that obtained in conventional liquid-phase phosphorylation (9, 10). Phosphorylation screening has several advantages over other methods including conventional substrate purification, determination of consensus peptide sequences, or interaction screening:

(a) Even scarce proteins can be detected as long as they are good substrates.

(b) Unlike peptide library screening methods, naturally existing proteins can be directly identified.

(c) Unlike interaction screening methods, substrates that do not form a stable complex with the protein kinase can be detected.

(d) Phosphorylation screening identifies only substrate proteins, whereas interaction screening is likely to detect also other kinase-interacting molecules such as subunits or regulatory proteins.

(e) Cloned cDNAs can be easily sequenced, and because they are produced as GST fusion proteins they can be readily purified for antibody preparation.

There are, however, some potential disadvantages and practical problems in the application of this technique:

(a) If the protein kinase of interest phosphorylates endogenous protein(s) derived from *E. coli* or λ phage, it is difficult to distinguish positive

plaques from negatives owing to high background. This problem, however, might be minimized if an appropriate affinity system in which the recombinant products are selectively retained on a membrane filter is available [e.g. a glutathione (GSH)-derivatized cellulose filter, see Section 8].

(b) A significant amount (e.g. microgram scale) of purified protein kinase is required in an active and soluble form.

(c) The screening may isolate not only physiological targets, but also proteins that are not physiological targets but, fortuitously, are good substrates, especially in the case of a protein kinase that has relatively low specificity.

(d) Protein kinases, or other proteins that autophosphorylate, may also score as positive in the screen. This problem can be reduced by pre-incubation of the filters with unlabelled ATP and is also minimized if the cDNA inserts are not too long. Using randomly primed cDNAs of ~1 kb in size for the library preparation might minimize problems with autophosphorylating protein kinases (see Section 3.3).

Prior to using the phosphorylation screening technique, these advantages and disadvantages must be fully considered to judge whether this approach is suitable for the protein kinase in question.

3. Construction of λGEX5 cDNA library

3.1 λGEX5 vector

For the new screening method, we modified the phage vector λgt11 to produce λGEX5 (5). The λgt11 vector has previously been successfully used for various expression screening strategies such as immunoscreening, nucleic acid–protein interaction (south-Western), and protein–protein interaction (far-Western) (11, 12). The λGEX5 vector contains a plasmid sequence between the two *Not*I sites, consisting of a ColE1 origin, the ampicillin resistance gene, and a GST gene followed by a small (0.43-kb) stuffer sequence (*Figure 1*). The nucleotide sequence of the plasmid region (pGEX-PUC-3T, see *Figure 4*) is available in the DDBJ/EMBL/Genbank database with the accession number AB014641. The *Sfi*I sites at the ends of the stuffer region are used for insertion of cDNA of up to 9 kb in size, which can be expressed as GST fusion proteins upon induction by isopropyl β-D-thiogalactopyranoside (IPTG). The λGEX5 vector system has the following advantages over the original λgt11 vector:

- Isolated clones can be rapidly converted into plasmid clones by excision rescue without purifying cDNA fragments.
- The rescued plasmids can be directly utilized not only for cDNA sequencing but also for expression of the GST fusion proteins for further characterization.

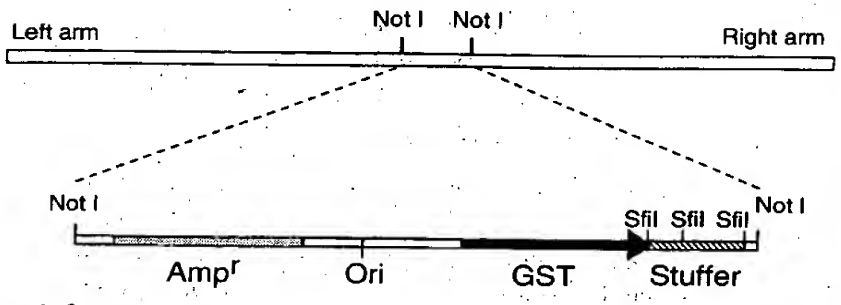


Figure 1. Structure of λGEX5 cloning vector. The plasmid region between two *NotI* sites is expanded in the lower part.

- GST, the N-terminal fusion partner of the recombinant product expressed by λGEX5, is highly soluble, easy to purify by GSH–agarose chromatography, and much smaller (27 kDa) than the β-galactosidase (114 kDa) component of fusion proteins expressed by λgt11 (12).

These advantages enable rapid isolation, subcloning, and characterization of a large number of positive clones at the same time.

3.2 Preparation of vector arms

λGEX5 phage DNA is purified from large quantities (10^{12}–10^{13} pfu) of phage lysate by established methods (13, 14). The λGEX5 phage contains an amber

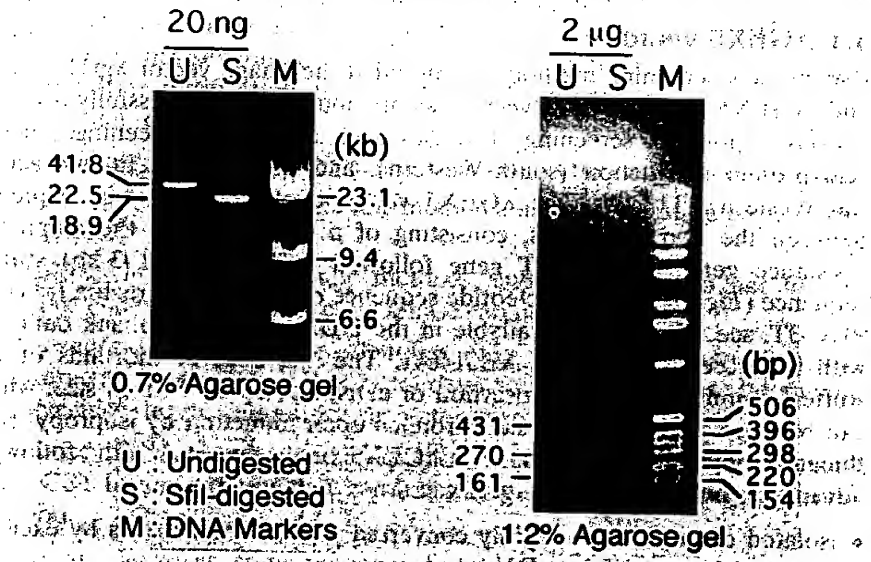


Figure 2. Confirmation of complete *SfiI* digestion of λGEX5 DNA for preparation of vector arms. Undigested or *SfiI*-digested λGEX5 DNA was analysed by electrophoresis on 0.7% (left) and 1.2% (right) agarose gels. The left arm (22.5 kb) and right arm (18.9 kb) are not clearly separated. Note that no DNA fragment is visible at the position of 431 bp.

mutation (*Sam*100) and should be propagated using *E. coli* BB4 (15) as the host strain. Other strains, such as Y1090 (12) and its derivatives, may not be suitable. If it is hard to obtain a high titre phage lysate by liquid culture a large-scale plate lysate (e.g. using 20 × 150 mm agarose plates) should be used instead (13, 14).

Vector arms for construction of a cDNA library are prepared by *Sfi*I digestion of the λGEX5 DNA. Dephosphorylation of the cleaved sites is not necessary because the 3′-protruding, single-stranded termini (3′-CGT) of the arms are not compatible with each other. The existence of an additional *Sfi*I site in the stuffer sequence helps to check whether or not the *Sfi*I digestion is complete (*Figure 2*). After confirming that the digestion is complete, the left arm (22.5 kb) and right arm (18.9 kb) are purified together by sucrose density gradient centrifugation or by preparative agarose gel electrophoresis (13, 14).

Protocol 1. Preparation of λGEX5 vector arms

Reagents

- λGEX5 DNA[a]
- TE buffer (10 mM Tris–HCl, pH 8.0 or 7.5, 1 mM EDTA)
- SDG buffer containing 10% or 40% sucrose [10 mM Tris–HCl, pH 8.0, 0.9 M NaCl, 5 mM EDTA, 10 or 40% (w/v) sucrose]

Method

1. Digest 80 μg of λGEX5 DNA with *Sfi*I.
2. Analyse two aliquots (20 ng and 2 μg) of the digested DNA by 0.7% and 1.2% agarose gel electrophoresis, respectively. It is important to confirm that no DNA is visible at the position of the undigested λGEX5 DNA on the 0.7% gel. On the 1.2% gel, 270 bp and 161 bp bands, but not a 431 bp band, should be visible. If the 431 bp band is detectable, it means that the *Sfi*I digestion is incomplete, which will cause a high ratio of empty phage clones in the library.
3. Extract the digested DNA once with phenol/chloroform, and once with chloroform, and recover the DNA by ethanol precipitation.
4. Dissolve the DNA in 0.4 ml of TE buffer pH 8.0, add 4 μl of 1 M $MgCl_2$, and incubate for 1 h at 42°C to allow the cohesive termini of the arms to anneal.
5. Prepare a 10–40% linear sucrose gradient (12 ml) in two centrifuge tubes [Beckman SW41.Ti rotor (or equivalent)].
6. Load 0.2 ml of the annealed DNA (40 μg) on to each gradient and centrifuge at 180 000 g (at r_{max}) for 16 h at 15°C in a Beckman SW41.Ti rotor (or equivalent).
7. Collect 10 drops (0.6–0.7 ml) of fractions through a 20G needle from the bottom of the centrifuge tube.

Protocol 1. *Continued*

8. Remove 10 μl of each fraction and analyse by 0.7% agarose gel electrophoresis.
9. Pool the fractions that contain the annealed and unannealed arms. Dilute the pooled fractions with an equal volume of TE buffer pH 7.5 and recover the DNA by ethanol precipitation.

*The λGEX5 vector is available on request through the DNA bank laboratory at Tsukuba Life Science Center, the Institute of Physical and Chemical Research (RIKEN), Tsukuba 305-0074, Japan. The ID number of the λGEX5 vector is RDB1911.

3.3 Preparation of *Sfi*I adaptor-ligated cDNA

Although synthesis of double-stranded cDNA is an important step for cDNA cloning, a detailed description of this step is beyond the scope of this chapter. Double-stranded cDNA can be synthesized from poly(A)$^+$RNA using an oligo(dT) primer or random hexanucleotide primers by a general method for cDNA synthesis (13, 14) or using an appropriate kit. Oligo(dT)-primed cDNA libraries are prone to be rich in clones that contains only the C-terminal part of a protein. On the other hand, randomly primed cDNA libraries evenly cover every part of a protein in principle, although they often contain a high percentage of cDNAs for ribosomal RNAs. At the final stage of cDNA synthesis, the ends of the double-stranded cDNA should be blunted for adaptor ligation. Two 5′-phosphorylated oligonucleotides are annealed to form an adaptor, which has a blunt end and a non-palindromic, single-stranded overhang (3′-ACG) compatible with the cloning sites of the *Sfi*I-digested λGEX5 arms. After adaptor ligation, the cDNA is separated by agarose gel electrophoresis, and the cDNA whose size is larger than ~0.8 kb in length is recovered by electroelution. If the cDNA is synthesized with random primers, it may be a good idea to fractionate only cDNA of ~1 kb in size, in order to minimize problems with autophosphorylating protein kinases. A large amount of excess oligonucleotide can be removed through the size fractionation step.

Protocol 2. Adaptor ligation and size fractionation of double-stranded cDNA

Reagents

- 5′-phosphorylated oligonucleotides [12 mer: 5′-pd(CCAGCACCTGCA)-3′; 9 mer: 5′-pd(AGGTGCTGG)-3′]
- 10 × ligase buffer (0.5 M Tris-HCl, pH 7.5, 0.1 M $MgCl_2$, 0.1 M DTT, 10 mM ATP)

Method

1. Make up an oligonucleotide mixture containing 2.4 and 1.8 μg of the 12-mer and 9-mer oligonucleotides, respectively, in 20 μl of 10 mM $MgCl_2$.

2. Incubate the mixture at 80°C for 2 min, allow to cool slowly to room temperature over a period of about 60 min, and then chill on ice.
3. Mix in the following order:

   55 μl of H_2O
   20 μl of the annealed oligonucleotides
   10 μl (1–2 μg) of double-stranded cDNA
   10 μl of 10 × ligase buffer
   1 μl of 10 units/μl of T4 polynucleotide kinase
   4 μl of 400 units/μl T4 DNA ligase
4. Incubate the mixture (100 μl) for 6–14 h at 16°C. Recover the DNA by phenol/chloroform extraction and ethanol precipitation.
5. Dissolve the adaptor-ligated cDNA in 20–50 μl of TE buffer pH 8.0, and separate the DNA fragments by preparative electrophoresis on a 1% agarose gel.
6. Locate the region of DNA fragments larger than 0.8 kb in size by using appropriate DNA size markers, and recover the size-fractionated DNA from the agarose gel by electroelution.

3.4 Ligation and *in vitro* packaging

The adaptor-ligated and size-fractionated cDNA molecules can now be ligated to the *Sfi*I-digested vector arms, followed by packaging of the ligated DNA into bacteriophage λ particles using an *in vitro* packaging reaction. We usually use commercially available kits for this ligation (e.g. DNA ligation kit Version 1, PanVera/Takara) and *in vitro* packaging (Stratagene Gigapack Gold, or equivalent) according to the manufacturer's instructions. It is desirable to carry out pilot ligations to optimize the ratio of cDNA to vector arms for efficient production of recombinant phages. Also, a control ligation without cDNA is essential to check the background level of empty phage. The titre of the *in vitro*-packaged phage is determined on *E. coli* BB4. The pfu in the original, packaged phage is the number of independent clones in the library. If necessary, the cDNA library can be amplified once by propagating phages in *E. coli* BB4 on agar plates (13, 14).

3.5 Construction of a positive control phage for screening

In phosphorylation screening, conditions for screening must be determined experimentally for each protein kinase. For this purpose, ideally one needs to have a positive control phage that expresses the GST fusion protein of an appropriate substrate protein (or peptide) for the protein kinase being used for the screening (assuming that one is known). The positive control phage can be constructed by ligating the *Sfi*I-digested λGEX5 arms to an appropri-

ate substrate cDNA produced by PCR with *Sfi*I site-containing primers. For the screening of ERK1 MAP kinase substrates, we constructed a λGEX5 recombinant encoding the C-terminal region of a transcription factor Elk-1 (λGEX-Elk-C), which contains multiple ERK phosphorylation sites (5, 16). Similarly, a GST fusion protein of the C-terminal domain of the retinoblastoma protein (GST–Rb) worked well as a positive control in a screen for cyclin E/Cdk2 substrates (6).

4. Phosphorylation screening

In this section we describe protocols for solid-phase phosphorylation, using as an example substrate screening with ERK1 MAP kinase (5). Points to be considered for application to other protein kinases will also be discussed.

4.1 Preparation of protein kinase for phosphorylation screening

The protein kinase used for phosphorylation screening must be soluble, active, and sufficiently pure. Care should be taken to avoid the presence of significant amounts of other protein kinases. Large-scale production and purification of recombinant protein kinase would be the best for this purpose. To obtain a large amount of activated ERK1 MAP kinase, we produced recombinant human ERK1 using a baculovirus expression system, in which insect Sf9 cells were co-infected with three recombinant baculoviruses encoding v-Ras, c-Raf-1, and ERK1 as described (5, 17). Similar co-expression strategies utilizing a eukaryotic expression system would also be effective for large-scale production of a protein kinase that requires a subunit protein or an activating modification such as phosphorylation. For example, active cyclin-Cdk complexes can be efficiently produced by co-expression of a hexahistidine-tagged cyclin and its partner Cdk in Sf9 cells, and easily purified using nickel-affinity column chromatography (6).

4.2 Preparation of plaque-immobilized filters

Procedures in this step are essentially the same as those used for expression screens using the λgt11 phage system (11–14). A λGEX5 cDNA library is plated on agar plates, and expression of the encoded GST fusion proteins is induced by overlaying IPTG-containing nitrocellulose filters on to the plaques. Plating density should be in the range of $1.5–3 \times 10^4$ plaques per 150 mm agar plate. High density plating ($>5 \times 10^4$ plaques per plate) may result in weak and small signals that are indistinguishable from false-positive signals.

Protocol 3. Plating out λGEX5 library and immobilization of plaques on filters

Reagents

- *E. coli* BB4 strain (Stratagene)
- SM (50 mM Tris–HCl, pH 7.5, 100 mM NaCl, 10 mM $MgSO_4$ and 2% gelatin)
- NZCYM medium (1% NZ amine, 0.1% casamino acids, 0.5% bacto-yeast extract, 1% NaCl, 10 mM $MgSO_4$; adjust the pH to 7.5 with 1 N NaOH)
- agar plates (60–80 ml of 1.5% agar in NZCYM medium per 150 mm plate)
- TB medium (1% bacto-tryptone, 0.5% NaCl)
- top agarose (0.7% agarose in NZCYM medium)
- nitrocellulose filters [137 mm in diameter, Schleicher & Schuell BA85 (0.45 μm) or equivalent]
- isopropyl β-D-thiogalactopyranoside (IPTG, 10 mM)

Method

1. Pick up a single colony of BB4 and grow cells in TB medium containing 0.2% maltose, 10 mM $MgSO_4$, and 12.5 μg/ml tetracycline[a] at 30°C overnight. Centrifuge the cells at 1500 *g* for 10 min and resuspend in 10 mM $MgSO_4$ at a density of A_{600} = 2.0. The plating bacteria can be stored at 4°C for up to a week.
2. Prepare 1.5–2 × 10^5 pfu/ml of λGEX5 library phage in SM.
3. Mix 0.1 ml (i.e. 1.5–2 × 10^4 pfu) of the phage with 0.5 ml of plating bacteria and incubate for 15 min at 37°C.
4. Add 8 ml of molten (50°C) top agarose, mix well, and pour on to 1.5% agar plates pre-warmed at 37°C. Leave the plates at room temperature for 15 min to harden the top agarose.
5. Incubate at 42°C for 3–4 h.
6. Soak nitrocellulose filters with 10 mM IPTG and remove the excess liquid by laying on Whatman 3MM paper.
7. Overlay the plates with the IPTG-impregnated nitrocellulose filters. Do not allow the plates to cool.
8. Incubate at 37°C for another 6–10 h.

[a] As BB4 contains F′ factor encoding *lacI*^{q} and *tet* genes, the original strain should be maintained in the presence of tetracycline.

4.3 Solid-phase phosphorylation

The next step is incubation of the plaque-immobilized filters in a reaction buffer containing protein kinase and [γ-^{32}P]ATP. Prior to the phosphorylation step, the filters are incubated in the presence of unlabelled ATP. The purpose of this pre-incubation step is to reduce the frequency of isolating clones whose products have an autophosphorylating or ATP-binding activity, as discussed in Section 2. *Protocol 4* describes an example of this step with

conditions that were optimized to carry out substrate screening for ERK1 MAP kinase (5).

Protocol 4. Solid-phase phosphorylation with ERK1 MAP kinase

Equipment and reagents

- rotating platform
- blocking solution (20 mM Tris–HCl, pH 8.0, 150 mM NaCl, 3% BSA, 1% Triton X-100)
- Triton wash buffer (20 mM Tris–HCl, pH 7.5, 150 mM NaCl, 10 mM EDTA, 1 mM EGTA, 0.5% Triton X-100, 1 mM DTT, 0.2 mM PMSF)
- MAPK reaction buffer (20 mM Hepes–NaOH, pH 7.5, 10 mM $MgCl_2$, 50 mM Na_3VO_4, 5 mM β-glycerophosphate, 5 mM NaF, 2 mM DTT, 0.1% Triton X-100)
- MAPK wash buffer (20 mM Tris–HCl, pH 7.5, 150 mM NaCl, 10 mM EDTA, 1 mM EGTA, 20 mM NaF, 0.1% Triton X-100)

Method

1. Cool the plates to room temperature. Mark the filters asymmetrically with a needle.
2. Carefully peel the filters off and immerse one by one into a large volume (at least 200 ml for 20 filters) of blocking solution. Throughout the following steps, filters should be kept wet and under gentle agitation (unless stated otherwise).
3. Agitate the filters slowly on a rotating platform for 60 min at room temperature.
4. Wash the filters three times for 20 min at room temperature in 200–300 ml of Triton wash buffer. Slightly more vigorous agitation is required to remove bacterial debris from the filters.
5. Wash the filters for 10 min at room temperature in 200 ml of MAPK reaction buffer.
6. Incubate the filters for 60 min at room temperature in 200 ml of MAPK reaction buffer containing 25–100 μM unlabelled ATP to mask proteins that have autophosphorylating and/or ATP-binding activities.
7. Wash the filters for 10 min in 200 ml of MAPK reaction buffer without ATP.
8. Incubate the filters for 60 min at room temperature (or 30°C if an air incubator is available) with gentle shaking in the MAPK reaction buffer containing 25 μM unlabelled ATP, 5 μCi/ml [γ-^{32}P]ATP, and 1 μg/ml purified human ERK1 MAP kinase. Use at least 2 ml of the solution per 137 mm filter. A tissue culture dish (e.g. Falcon 150 mm dish, #3025) is convenient for the incubation of up to 20 filters.
9. Wash the filters 6–7 times for 5–10 min at room temperature in 100–200 ml of MAPK wash buffer. Relatively vigorous agitation helps to reduce false-positive signals. Finally, wash the filters once with MAPK wash buffer without Triton X-100.

10. Dry the filters on paper towels, and arrange on a paper sheet for autoradiography.

4.4 Identification of cDNA clones encoding substrate candidates

The phosphorylated filters are exposed to X-ray film to identify positive plaques by autoradiography. The intensities of radioactive signals are generally quite variable from plaque to plaque, presumably because the recombinant proteins are expressed at different levels, and also because they are not all equally efficient substrates and are therefore phosphorylated to different extents. Therefore, for practical reasons, it is advisable to decide to pick up only a limited number (perhaps ~100) clones, in order of intensity, from the plates in the primary screening. It may often be difficult to discriminate truly phosphorylated plaques from false-positive spots. This problem can be overcome by making an additional autoradiogram of the same filter with much shorter exposure. Generally speaking, pinpoint signals with a sharp rim found in the short exposure are false-positives, whereas true-positive plaques give rather dull, fuzzier signals with a certain range of sizes. Therefore, most false-positive signals can be easily excluded by carefully comparing signals in the corresponding position in the two autoradiographs. If it is possible, it will also help to include a positive λGEX substrate control plate, since this will indicate the sort of signal intensity and morphology to expect for true positive plaques in the screen (see Section 4.5). Alternatively, primary screening may be carried out in duplicate (13). By comparing the duplicate filters, true plaques can be easily identified as reproducible signals. Since most plaques are in contact with one another in the primary screening, phage clones must be purified by secondary screening in which the phages are plated out more sparsely. This step also confirms whether the proteins expressed in the plaques identified in the primary screening are truly phosphorylated by the protein kinase.

Protocol 5. Identification and isolation of positive clones

Equipment and reagents

- as *Protocols 3* and *4*

Method

1. Expose the phosphorylated filters to X-ray films for 4–72 h at –70°C with an intensifying screen. Make at least two sets of autoradiograms with different exposure times, one with a long exposure and the other with a much shorter exposure.
2. Mark strong signals on the long exposure films and carefully observe

Protocol 5. *Continued*

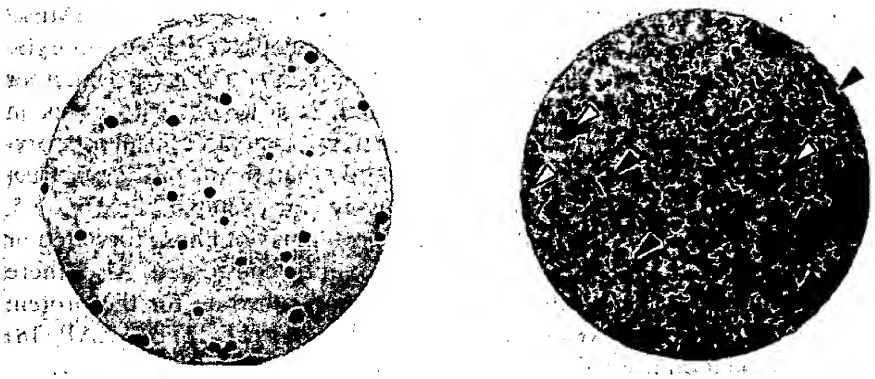
their corresponding signals on the short exposure films. Identify sharp, pinpoint signals on the short exposure films, which should be false-positives. Mark them again differentially on the long exposure films for exclusion.

3. Align the long exposure films with the filter sheets. Mark the positions of the asymmetric needle spots on the filters.
4. Identify the locations of positive plaques, and remove a relatively large agar plug (2–4 mm in diameter) from each position to be certain of recovering the positive clone.
5. Transfer the agar plug to 1 ml of SM containing 2 drops of chloroform, and incubate for 1–2 h at room temperature. Measure the titre of each phage stock.
6. Replate each phage stock on to a 90 mm agar plate at a low density (200–500 plaques per plate) and transfer the plaques to 85 mm nitrocellulose filters as described in *Protocol 3.*
7. Repeat solid-phase phosphorylation of the plaque filters as described in *Protocol 4.* Use 0.7 ml of the kinase reaction buffer per filter. Expose the phosphorylated filters to X-ray films.
8. Identify a single, well-isolated positive plaque (clone) on each plate. If it is difficult to obtain a single plaque, repeat steps 3–8 to purify a phage clone.

Figure 3 shows a typical result of solid-phase phosphorylation using ERK1 MAP kinase. In the control experiment with positive (λGEX-Elk-C) and negative (λGEX5) control phages, these two phages gave a clear contrast in intensity of radioactive signals. In the actual screening of a HeLa cDNA library, however, signals of various degrees of intensity are observed as discussed above.

4.5 Optimization of screening conditions using control phages

Since optimum conditions for phosphotransfer reactions are quite variable amongst protein kinases, screening conditions should be experimentally determined for each protein kinase being used. Components of reaction buffer to be considered would be pH, species and concentrations of monovalent and divalent cations, type of detergent, protease inhibitors, phosphatase inhibitors, other additives, and so forth, depending on the enzymological properties of the protein kinase. The presence of detergent (e.g. Triton X-100, Nonidet P-40, Tween 20, etc.) in the buffers is helpful to reduce false-positive signals, most of which seem to be derived from dust or bacterial debris. Thus, it is



λGEX-Elk-C : λGEX5 = 1:100 HeLa cDNA library

Figure 3. Autoradiogram of a solid-phase phosphorylation screen. Phage plaques (4000 pfu per 90 mm agar plate) were transferred to 85 mm nitrocellulose filters and then subjected to phosphorylation screening with ERK1 MAP kinase according to *Protocols 3* and *4*. Left: a 1:100 mixture of the positive control phage (λGEX-Elk-C) and a negative control phage (λGEX5). Right: screening of HeLa cDNA library. Filled triangles indicate positive plaques whereas open triangles indicate false-positives. Although these signals cannot be distinguished in this figure, signals indicated by open triangles showed up as pinpoint spots on a short exposure film (not shown).

desirable to include a detergent at least in the washing buffers unless the protein kinase is highly sensitive to it. There are no general rules for the concentration and specific radioactivity of [γ-^{32}P]ATP, concentration of protein kinase, reaction temperature, and duration. Therefore, the first trial should be performed in the reaction mixture that is usually used for the in-solution kinase assay of the protein kinase. In practice, however, a simple scaling up from test-tube (10–50 μl) to culture dish (10–50 ml) may not be easy, especially for the quantities of the protein kinase and radioactive ATP needed. Fortunately, a reduction in concentration of these components may be compensated to some extent by raising the specific activity of [γ-^{32}P]ATP (i.e. reducing the concentration of unlabelled ATP) and by changing the incubation time for phosphorylation and the exposure time of autoradiography. Optimization of these three parameters is particularly important to obtain a good signal-to-noise ratio in the autoradiogram.

Ideally, determination of screening conditions should be carried out using positive and negative control phages. If a substrate for the protein kinase in question is already known, the positive control phage can be constructed as described in Section 3.5. If not, reaction conditions may need to be checked while an actual screen is being performed with a cDNA library, assuming that positive clones exist. The λGEX5 vector phage can usually be used as a negative control phage, although this phage expresses a GST-fusion protein with an artificial C-terminal sequence derived from the stuffer-region (GST–

stuffer). If the stuffer-derived sequence (58 amino acids; DDBJ/EMBL/GenBank accession number AB014641) contains any potential phosphorylation sites for the protein kinase, it may be necessary to construct another control phage expressing only the GST region. It is desirable to check in advance whether the GST–fused substrate protein (or peptide) is phosphorylated by the protein kinase much more efficiently than the negative construct (GST–stuffer or GST itself) in a test-tube assay (for example, see *Figure 5*, lanes of GST and GST-Elk-C). If the fusion protein is not phosphorylated or is poorly phosphorylated, other constructs should be considered. Also, there is a possibility that GST itself may act as a good substrate for the protein kinase. In this case, other vector systems such as λgt11 (11, 12) or λZAP (18) may have to be used instead.

The control experiment can be performed using 90 mm agar plates on which either of the control phages and a series of mixtures of the two control phages are plated out at a density of 4–8 × 10^3 plaques per filter. The ratios of the positive and negative control phages in the mixtures should be, for example, 1:10, 1:100, and 1:1000. The autoradiogram of solid-phase phosphorylation of these filters helps to estimate the difference in intensity between positive and negative signals, the background level in radioactivity of plaques and *E. coli* lawn, and the frequency of appearance of false-positive signals. If the model experiment works well, the number of strong signals should change amongst the filters in direct proportion to the ratio of the positive phages in the mixtures (*Figure 3*).

5. Conversion of phage clones into plasmids

If positive plaques give clearly stronger signals than the remaining clones in the secondary screening, this means that the first step of the screening is successful. For further characterization, positive clones now have to be converted into plasmids that contain their cognate cDNA. *Protocol 6* describes a rapid plate

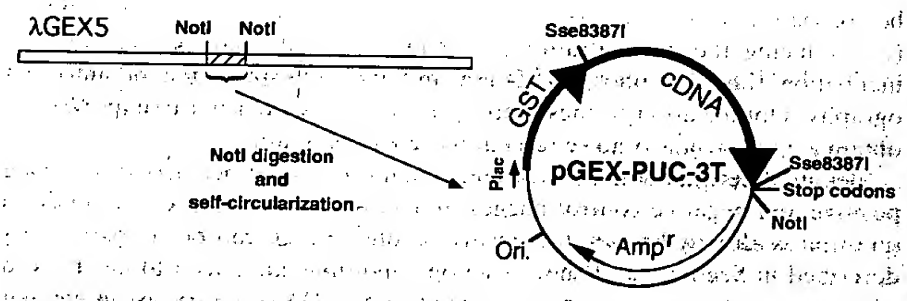


Figure 4. Schematic representation of excision rescue. After digestion of phage DNA with *Not*I, the plasmid region containing cDNA can be recovered by self-ligation. In the resulting pGEX-PUC-3T plasmid, the *Sfi*I sites used for cDNA cloning are not regenerated but *Sse*8387I sites are available to excise the cDNA fragment.

lysate method for preparing DNA from multiple (10–100) phage clones at the same time. The cDNA-containing plasmid (pGEX-PUC-3T) can be recovered by *Not*I digestion of the phage DNA followed by self-circularization, as described in *Protocol 7*. We recommend *E. coli* XL1-Blue (15) as the host strain for the transformation of this plasmid. A schematic representation of this process is shown in *Figure 4*.

Protocol 6. Rapid, small-scale preparation of λGEX5 phage DNA

Reagents

- λ diluent (10 mM Tris–HCl, pH 7.5, 10 mM $MgSO_4$)
- 50 × DNase/RNase mixture (0.1 mg/ml DNase I, 1 mg/ml RNase A; make up 1 × DNase/RNase mixture in λ diluent just before use)
- 35 mm agarose[a] plates in 6-well tissue culture plate (3 ml of 1.5% agarose in NZCYM medium per well)
- extraction buffer (10 mM Tris–HCl, pH 8.0, 0.1 M NaCl, 10 mM EDTA, 0.1% SDS)

Method

1. Using a pasteur pipette, pick a single plaque and place in 0.5 ml of SM containing one drop of chloroform. Incubate the suspension for 1–2 h at room temperature.
2. Mix 20 μl of the phage suspension with 30 μl of plating bacteria (BB4, A_{600} = 2.0) and incubate for 15 min at 37 °C.
3. Add 0.4 ml of molten (50 °C) top agarose[a] (0.7% agarose in NZCYM medium), and spread the bacterial suspension on the surface of 35 mm agarose plates in 6-well tissue culture plates.
4. Incubate at 37 °C overnight (10–16 h) to reach confluent lysis.
5. Directly add 0.65 ml of λ diluent containing 2 μg/ml DNase I and 20 μg/ml RNase A on to the surface of the top agarose, and incubate for 1.5–2 h at room temperature with constant, gentle shaking.
6. Transfer the phage suspension (usually 0.4–0.5 ml) to a microfuge tube, add 30 μl of chloroform, and vortex for 5–10 sec.
7. After centrifugation for 1 min, transfer 0.4 ml of the aqueous supernatant to a microfuge tube. Do not take any bacterial debris.
8. Add 0.4 ml of λ diluent containing 20% (w/v) polyethylene glycol 8000 and 2 M NaCl. Mix by vortexing, and incubate for 1 h on ice.
9. Centrifuge at 10 000 *g* for 10 min at 4 °C. Remove the supernatant, leaving 30–50 μl behind. Recentrifuge briefly to bring the liquid on the walls of the tube to the bottom, and remove the remaining supernatant.
10. Add 100 μl of extraction buffer, dissolve the phage pellet by vortexing, and incubate for 10 min at 68 °C.

Protocol 6. *Continued*

11. Extract the solution with 100 μl of phenol/chloroform, and recover the phage DNA by ethanol precipitation.
12. Dissolve the DNA in 100 μl of TE buffer pH 8.0.

[a] Agarose, not agar, should be used. Phage DNA prepared from agar plates is often resistant to restriction enzyme digestion.

Protocol 7. Excision rescue of cDNA-containing plasmid from λGEX5 clones

Equipment and reagents

- *Not*I reaction mixture [for 10 reactions, 45 μl of water, 10 μl of 10 × *Not*I buffer (NEB or equivalent) and 5 μl (5–25 units) of *Not*I; prepare the mixture on ice just before use]
- 2X YT medium (1.6% bacto-tryptone, 1% bacto-yeast extract, 0.5% NaCl)
- ligase reaction mixture [for 10 reactions, 40 μl of water, 7 μl of 10 × ligase buffer (*Protocol 2*), and 3 μl (300–1200 units) of T4 DNA ligase; prepare the mixture on ice just before use]
- XL1-Blue competent cells (Stratagene)

Method

1. Place 4 μl of each phage DNA solution from *Protocol 6* in a microfuge tube. Add 6 μl of a *Not*I reaction mixture, incubate at 37°C for 1 h, and then at 70°C for 20 min.
2. Centrifuge the tubes briefly, and then transfer 2 μl of the *Not*I-digested DNA into a well of a U-bottom, 96-well microtitre plate.
3. Add 5 μl of ligase reaction mixture. Mix by vortexing the plate, and incubate for 30 min at 16°C.
4. Add 20 μl of XL1-Blue competent cells directly to the ligated DNA, mix, and incubate on ice for 5 min. Heat for 90 sec at 42°C, add 200 μl of 2X YT medium, and incubate for 30 min at 37°C with gentle shaking.
5. Transfer an appropriate volume (200 μl per 90 mm plate or 50–100 μl per 35 mm plate in a 6-well plate) of the cell suspension on to an agar plate containing 100 μg/ml ampicillin.
6. Incubate overnight (>10 h) at 37°C. Most of the clones thus obtained usually contain plasmids of the expected structure (i.e. pGEX-PUC-3T containing a cDNA between the *Sse*8387I sites), but it may be a good idea to pick up two colonies from each plate for back-up.

6. Initial characterization of candidate clones

First of all, cDNA clones from positive plaques need to be checked with respect to the following criteria:

- Can the cDNA-encoded polypeptide really act as a substrate for the protein kinase?
- Does the polypeptide sequence represent at least a part of any naturally existing protein?

The former point can be examined by *in vitro* phosphorylation assay of each of the purified GST fusion proteins. The latter question arises since a polypeptide translated from a wrong reading frame or non-coding region of cDNA could be a good substrate by chance. Fortunately, clones that encode such artefactual products can be effectively excluded from the list of candidates by checking their product sizes on SDS–PAGE (see Section 6.1).

6.1 Purification and analysis of GST fusion proteins encoded by candidate clones

GST-fused recombinant proteins encoded by the rescued plasmids can be expressed in small bacterial cultures, purified by GSH–agarose affinity beads, and then tested for phosphorylation by the protein kinase. An SDS–PAGE analysis of the phosphorylated GST fusion proteins offers information about their sizes and extent of phosphorylation.

Protocol 8. Analysis of GST fusion proteins produced by candidate clones

Equipment and reagents

- sonicator with a microtip
- TLB (50 mM Hepes–NaOH, pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton-X100, 1 mM EGTA, 1 mM DTT, 1 mM PMSF)
- GSH–agarose beads (Sigma, or equivalent)
- lysozyme solution (5 mg/ml in 0.1 M Tris–HCl, pH 8.0)
- 2 × SB (0.1 M Tris–HCl, pH 6.8, 4% SDS, 20% glycerol, 40% 2-mercaptoethanol, 0.04% bromophenol blue)

Method

1. Pick a single colony and inoculate 3 ml of 2 × YT medium containing 0.1% glucose and 100 μg/ml ampicillin. Incubate overnight (12–18 h) at 37°C with vigorous shaking.
2. Place 0.15 ml of the overnight culture into a microfuge tube containing 50 μl of 60% sterile glycerol. Mix well and store at –70°C as a glycerol stock.
3. Place 1.35 ml of the overnight culture into a microfuge tube and prepare plasmid DNA by a miniprep method suitable for DNA sequencing.
4. Add 7.5 μl of 100 mM IPTG to the remaining culture (1.5 ml), and incubate for 3–4 h at 37°C.
5. Transfer the culture into a microfuge tube, centrifuge at 10 000 *g* for

Protocol 8. *Continued*

30 sec at 4 °C. Remove the supernatant and resuspended the bacterial pellet in 100 μl of ice-cold TE (pH 7.5).

6. Add 10 μl of freshly prepared lysozyme solution, mix, and leave on ice for 5 min.
7. Add 0.5 ml of TLB and lyse the cells by sonication. Centrifuge the tube at 12 000 *g* for 15 min at 4 °C and transfer the supernatant to a fresh tube.
8. Add 100 μl of 50% (v/v) GSH–agarose beads and incubate for 1 h at 4 °C with rocking.
9. Wash the beads twice with 0.5 ml of ice-cold TLB and once with 1 ml of MAPK reaction buffer (see *Protocol 4*). Add to the washed beads 50 μl of the same buffer and mix by vortexing.
10. Transfer 20 μl of the suspension containing about 10 μl of the beads into a fresh tube and add 5 μl of MAPK reaction buffer containing 250 μM unlabelled ATP, 0.5 μCi of [γ-^{32}P]ATP, and 10–50 ng of ERK1 MAP kinase. Incubate the reaction mixture (25 μl) for 30–60 min at 30 °C with occasional agitation.
11. Add 25 μl of 2 × SB, boil for 3–5 min, and then analyse 10–30 μl of the samples by SDS–PAGE on a 12.5% polyacrylamide gel.
12. Stain the gel with Coomassie blue, dry and expose to an X-ray film for autoradiography.

In our screen for ERK1 substrates, we screened about 3×10^5 independent clones of a cDNA library prepared from HeLa cells, and obtained 120 positive clones from the secondary screening (5). *Figure 5* shows a typical example of *in vitro* phosphorylation analysed by SDS–PAGE. Although almost all of the recombinant proteins were phosphorylated by ERK1 MAP kinase *in vitro*, more than half of them were GST proteins with a very short fusion partner (S5, S6, S8, S17, and S19). Sequencing analysis revealed that most of the extremely small fusion partners were artefactual products derived from out-of-frame ligation of cDNAs. Therefore, out of 120 clones, we selected 32 clones that expressed GST fusion proteins larger than 32 kDa in total size (i.e. with a fusion partner of >5 kDa) for further characterization (5). Every recombinant product of these 32 clones was a good *in vitro* substrate for ERK1, and is likely to represent some part of a naturally existing protein (see Section 6.2). Generalizing from this case, it is advisable to select only clones that produce a GST fusion protein whose size is larger than, for example, 32 kDa, which means that clones encoding cDNA-derived polypeptides of <5 kDa should be excluded from the substrate candidates. Although the frequency of appearance of such artefactual recombinants may depend on

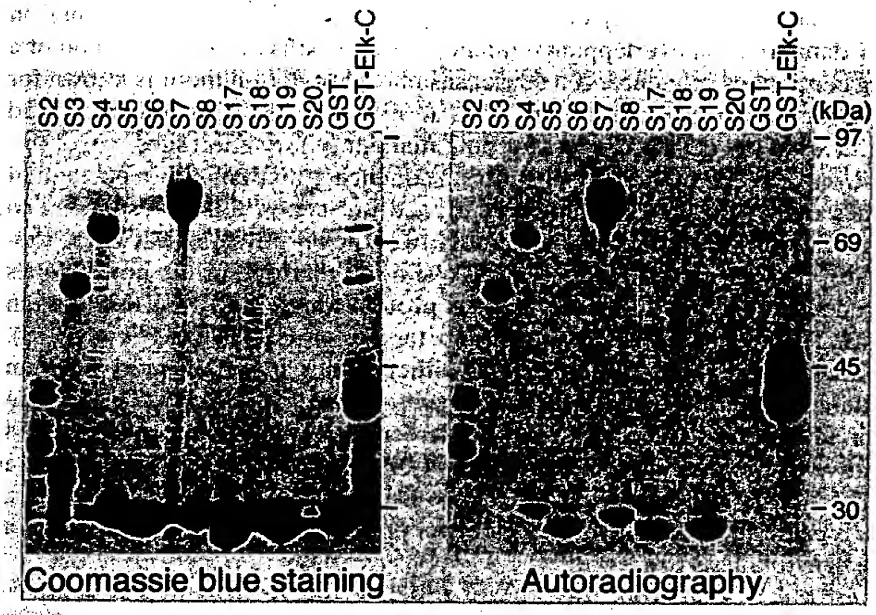


Figure 5. *In vitro* phosphorylation analysis of positive clones in a screen for ERK1 substrates. Positive phage clones (S2–S8 and S17–S20) obtained in the screening were converted into plasmid clones, and then their GST fusion proteins were expressed and purified using GSH–agarose. The resulting GST fusion proteins together with GST and GST–Elk-C were phosphorylated by ERK1 MAP kinase in the presence of [γ-^{32}P]ATP and analysed by SDS–PAGE. Left: Coomassie blue staining of the gel. Right: autoradiogram of the same gel. Note that GST itself was not phosphorylated at all, whereas the artefactual recombinants (S5, S6, S8, S17, and S19) with similar small sizes and quantities were relatively weakly phosphorylated.

the protein kinase, it is probably desirable to exclude artefacts by checking the size of each GST fusion protein prior to sequencing.

6.2 Sequencing analysis

The 5′- and 3′-nucleotide sequences of the candidate clones can be determined using oligonucleotide primers designed to anneal just outside the cloning sites. Amino acid sequences of the fusion proteins can be easily deduced from the 5′-nucleotide sequence. The nucleotide and amino acid sequences should be subjected to identity/homology searches using appropriate programs such as BLAST (19). If a particular cDNA encodes a polypeptide sequence that is identical or similar to that of any known protein, this information may offer clues with regard to further analysis. Some of them may be very plausible candidates for physiological targets, whereas some others may be most unlikely (e.g. secreted proteins). The nucleotide sequences of some clones may be found in an EST (expressed sequence tag) database, which may offer

additional nucleotide sequence missing in the partial cDNA. 'Walking' in EST databases via overlapping sequences often results in the elucidation of a long contiguous sequence. If a consensus phosphorylation motif is known for the protein kinase in question, then it is worth scanning the predicted sequence of the cDNA product for potential phosphorylation sites.

In our screen for ERK1 substrates, sequencing analysis revealed that 14 clones out of the 32 candidates mentioned in Section 6.1 corresponded to fragments of known proteins, including two known physiological ERK substrates, $p90^{RSK2}$ and c-Myc (5). We also showed that one of the novel clones encoded a new MAP kinase-activated protein kinase and that this protein kinase, named MNK1, was an *in vivo* target for ERK1 MAP kinase (5). Although we have not characterized other clones in detail, most of them contained potential MAP kinase recognition sites, Ser–Pro or Thr–Pro, in their amino acid sequences. In another screen for cyclin E/Cdk2 substrates, 54 positive clones were identified from 1.0×10^6 independent clones of the HeLa cDNA library. At least two of them are likely to be *in vivo* Cdk substrates, one is caldesmon and the other is a novel protein, named PRC1, that plays a role in cytokinesis (6).

7. Analysis of physiological function of the identified substrates

If a large number of clones are isolated in a screen, it may be necessary to decide which clone(s) are to be further characterized. In the case of clones showing identity or homology to some known sequence, this information can help to inform this decision. In the case of novel proteins, however, it is not easy to choose one out of the candidates. Therefore, some additional experiments need to be done to guide selection. A simple experiment would be to test whether each GST–fused substrate protein can bind the protein kinase by an *in vitro* binding assay. If a substrate physically associates with the kinase to form a complex, it may be worth analysing further. Alternatively, analysis of expression patterns of the candidates by Northern or *in situ* hybridization may be helpful. If a clone shows a characteristic expression pattern such as cell type specificity, cell cycle dependence, developmental regulation, or inducible expression, this may suggest some physiological relationship between the kinase and substrate.

Finally, it is important to remember that a protein identified by phosphorylation screening remains simply an '*in vitro*' substrate until proven to be a physiological target of the protein kinase used, and this issue can be clarified only by *in vivo* analysis. It is important to determine whether the protein is phosphorylated under conditions where the protein kinase in question is known to be activated. The final goal is, of course, to elucidate the function of the newly identified substrates and the physiological significance of their

phosphorylation. Although there is no general strategy for this purpose, isolation and sequence determination of a full-length cDNA clone, as well as production of antibodies, are essential steps. Various experiments such as overexpression or ectopic expression, decreasing expression by antisense methods, or inhibition of protein function by antibody microinjection, may suggest possible roles for the protein *in vivo*. Also, it may be informative to examine the subcellular localization and kinetics of synthesis and degradation of the protein. Furthermore, determination of the sites phosphorylated *in vitro* and *in vivo*, and mutation of these sites, would help to elucidate the biological significance of phosphorylation in regulating the function of the protein kinase target.

8. Application to other protein kinases

The phosphorylation screening approach is, in principle, generally applicable to all protein kinases. A major problem that may arise in some cases, as discussed in Section 2, is a high background caused by strong phosphorylation of an endogenous protein(s) derived from *E. coli* or λ phage. The severity of this background noise may depend on the substrate specificity of the protein kinase used. Although we did not encounter this problem in the screens for ERK1 MAP kinase and cyclin E/Cdk2 substrates, which are both proline-directed protein kinases, a preliminary experiment with the catalytic subunit of cAMP-dependent protein kinase showed indistinguishable signals between positive and negative plaques. The high background problem may be overcome if it is possible to develop a GSH-derivatized filter on to which the GST-fused recombinants can selectively be immobilized. The use of this affinity filter system should, in principle, reduce the background problem caused by bacterial proteins. No such GSH filter is yet commercially available, but in preliminary experiments we have coupled GSH to cellulose filters and have found that the λGEX5/GSH–cellulose filter system seemed to work well for ERK1. However, further development is required for practical, large-scale usage.

Another modification for reducing the problem of a high background may be possible. Some protein kinases prove to form a relatively stable complex with their substrates through a domain distinct from the phosphorylation site, as is the case for JNK and c-Jun (20–22), and for certain MAP kinases and MAP kinase kinases (23, 24). As discussed in Section 7, this phenomenon might be of use to select some clones out of a great number of substrate candidates in the phosphorylation screening. Carrying this idea one step further, a modified screening method may be applicable to some protein kinases. In this alternative protocol, plaque filters are first pre-incubated with the protein kinase in the absence of ATP to allow the kinase to bind recombinant substrates. Then the filters are briefly washed to remove the excess, unbound protein kinase (different stringencies of washing could usefully be tried), and

subsequently incubated in the presence of [γ-^{32}P]ATP to allow phosphotransfer to occur within the kinase–substrate complex. We have not tested this alternative method, but it may reduce background signals derived from non-specific phosphorylation, and would be a way of identifying substrates that have a high affinity binding site. In fact, phosphorylation screening may work well because, with the relatively low protein kinase concentration used in the screen, only high affinity substrates are phosphorylated efficiently.

Application of phosphorylation screening to protein–tyrosine kinases may be bedeviled by the high background problem, because in general protein–tyrosine kinases are rather non-specific *in vitro*. The use of anti-phosphotyrosine antibodies to screen the phosphorylated λgt11 cDNA expression library, selecting only the strongest signals, has proved to be a successful way of identifying Src protein–tyrosine kinase substrates (7). An alternative solution would be to use an additional far-Western selection step utilizing a phosphotyrosine-binding domain. For example, plaque filters are first incubated with a protein–tyrosine kinase in the presence of unlabelled ATP. After washing, the filters are next incubated with an epitope-tagged or radiolabelled protein containing a particular SH2 or PTB (phosphotyrosine-binding) domain to allow them to bind, and then positive plaques are visualized by immunodetection or autoradiography. This protocol can be a substrate screen based not only on the phosphorylating specificity of the protein kinase, but also on the binding specificity of the phosphotyrosine-binding modules used. In fact, Kavanaugh and co-workers showed that a screen of a λgt11 cDNA library with solid-phase phosphorylation by platelet-derived growth factor receptor tyrosine kinase, followed by binding of ^{32}P-labelled Shc PTB domain, resulted in cloning of a c-ErbB2 (Neu) cDNA, which contained tyrosine phosphorylation-dependent PTB-binding sites (25). Similar 'double selection strategies' utilizing phosphorylation-dependent protein–protein interactions might be applicable to some protein–serine/threonine kinases, in combination with a protein module probe such as 14-3-3 proteins (26), Pin1 (27, 28), and the MPM-2 monoclonal antibody (29), each of which is shown to specifically interact with a phosphoamino acid-containing peptide motif.

We expect that the phosphorylation screening technique will be applicable, with further refinement, to other protein kinases including protein–tyrosine kinases and receptor-type protein kinases.

Acknowledgements

We are deeply appreciative of the support and advice of our colleagues, especially Wei Jiang for helpful discussions during the development of this method

References

1. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H., and Cantley, L. C. (1994). *Curr. Biol.*, **4**, 973.
2. Zhao, J., Dynlacht, B., Imai, T., Hori, T., and Harlow, E. (1998). *Genes Dev.*, **12**, 456.
3. Yang, X., Hubbard, E. J., and Carlson, M. (1992). *Science*, **257**, 680.
4. Waskiewicz, A., Flynn, A., Proud, C. G., and Cooper, J. A. (1997). *EMBO J.*, **16**, 1909.
5. Fukunaga, R. and Hunter, T. (1997). *EMBO J.*, **16**, 1921.
6. Jiang, W., Jimenez, G., Wells, N.J., Hope, T.J., Wahl, G.M., Hunter, T., and Fukunaga, R. (1998). *Mol. Cell*, **2**, 877.
7. Lock, P., Abram, C. L., Gibson, T., and Courtneidge, S. A. (1998). *EMBO J.*, **17**, 4346.
8. Smith, D. B. (1993). *Meth. Mol. Cell. Biol.*, **4**, 220.
9. Valtorta, F., Schiebler, W., Jahn, R., Ceccarelli, B., and Greengard, P. (1986). *Anal. Biochem.*, **158**, 130.
10. Glover, C. V. C. and Allis, C. D. (1991). In *Methods in enzymology* (ed. T. Hunter and B. M. Sefton), Vol. 200, p. 85. Academic Press, San Diego.
11. Young, R. A. and Davis, R. W. (1983). *Proc. Natl. Acad. Sci. USA*, **80**, 1194.
12. Huynh, T. V., Young, R. A., and Davis, R. W. (1985). In *DNA cloning, volume I: a practical approach* (ed. D. M. Glover), p. 49. IRL Press, Oxford.
13. Sambrook, I., Fritsch, E. F., and Maniatis, T. (ed.) (1989). *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, NY.
14. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (ed.) (1997). *Current protocols in molecular biology* (Supplement 37). John Wiley & Sons, Inc., NY.
15. Bullock, W. O., Fernandez, J. M., and Short, J. M. (1987). *BioTechniques*, **5**, 376.
16. Marais, R., Wynne, J., and Treisman, R. (1993). *Cell*, **73**, 381.
17. Williams, N. G., Paradis, H., Agarwal, S., Charest, D. L., Pelech, S. L., and Roberts, T. M. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 5772.
18. Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. (1988). *Nucl. Acids Res.*, **16**, 7583.
19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). *J. Mol. Biol.*, **215**, 403.
20. Hibi, M., Lin, A., Smeal, T., Minden, A., and Karin, M. (1993). *Genes Dev.*, **7**, 2135.
21. Kallunki, T., Su, B., Tsigelny, I., Sluss, H. K., Dérijard, B., Moore, G., Davis, R., and Karin, M. (1994). *Genes Dev.*, **8**, 2996.
22. Gupta, S., Barrett, T., Whitmarsh, A. J., Cavanaugh, J., Sluss, H. K., Dérijard, B., and Davis, R. J. (1996). *EMBO J.*, **15**, 2760.
23. Bardwell, L., Cook, J. G., Chang, E. C., Cairns, B. R., and Thorner, J. (1996). *Mol. Cell. Biol.*, **16**, 3637.
24. Fukuda, M., Gotoh, Y., and Nishida, E. (1997). *EMBO J.*, **16**, 1901.
25. Kavanaugh, W. M., Turck, C. W., and Williams, L. T. (1995). *Science*, **268**, 1177.
26. Muslin, A. J., Tanner, J. W., Allen, P. M., and Shaw, A. S. (1996). *Cell*, **84**, 889.
27. Ranganathan, R., Lu, K. P., Hunter, T., and Noel, J. P. (1997). *Cell*, **89**, 875.
28. Lu, P.J., Zhou, X.Z., Shen, M., and Lu, K.P. (1999). *Science*, **283**, 1325.
29. Davis, F. M., Tsao, T. Y., Fowler, S. K., and Rao, P. N. (1983). *Proc. Natl. Acad. Sci. USA*, **80**, 2926.